STUDIES IN THEORETICAL PSYCHOLINGUISTICS

The Acquisition of Verbs and their Grammar

The Effect of Particular Languages

NATALIA GAGARINA AND INSA GÜLZOW (EDS.)

Springer

THE ACQUISITION OF VERBS AND THEIR GRAMMAR: THE EFFECT OF PARTICULAR LANGUAGES

STUDIES IN THEORETICAL PSYCHOLINGUISTICS

VOLUME 33

The titles published in this series are listed on www.springer.com.

THE ACQUISITION OF VERBS AND THEIR GRAMMAR: THE EFFECT OF PARTICULAR LANGUAGES

Edited by

NATALIA GAGARINA
Zentrum für Allgemeine Sprachwissenschaft, Berlin, Germany

and

INSA GÜLZOW
Zentrum für Allgemeine Sprachwissenschaft, Berlin, Germany

Springer

A C.I.P. Catalogue record for this book is available from the Library of Congress.

ISBN 978-1-4020-4336-9 (PB)
ISBN 978-1-4020-4334-5 (HB)
ISBN 978-1-4020-4335-2 (e-book)

Published by Springer,
P.O. Box 17, 3300 AA Dordrecht, The Netherlands.
www.springer.com

Printed on acid-free paper

TABLE OF CONTENTS

NATALIA GAGARINA AND INSA GÜLZOW

INTRODUCTION

The present volume contains selected papers from the conference on the acquisition of verb grammar and verb arguments held at the *Research Center for General Linguistics* (ZAS), Berlin in November, 2001. The main aim of this conference was to provide a forum for researchers studying the (language-specific) acquisition of verbs and aspects of their grammar, such as inflection, finiteness, argument structure, etc. Of particular interest were the developmental processes at the interface of related components of verb grammar, such as lexicon and morphology, semantics and syntax, etc. Elucidating these issues should enhance our understanding of the development of particular verbal systems and of specific languages impact on this development.

The central role of the verb within a sentence has been acknowledged by most grammatical traditions and it is nowadays accepted by the two most prominent 'poles' in acquisition studies: the UG-based and the usage-based acquisition theories. These two psycholinguistic theories reflect two basic and opposed approaches to language. The first theory represents the most elaborated formalist approach (and provides corresponding accounts for language acquisition) while the second denies any principled difference between competence and performance and takes into consideration functions of the target language units the child has to acquire (cf. Newmeyer 2003 on the knowledge and use of language). Usage-based approaches may be generally characterized as emphasizing the crucial role of children's linguistic experience (in particular, frequency, statistical patterns, etc.) on the development of grammar. Linguistic competence in these theories is seen as a dynamic system which is continually shaped by linguistic usage events. The role of rule-based linguistic representations at the early stages of language development is downplayed in favor of constraint-based systems whose structural properties actually emerge from usage (cf. among others Barlow 2000, Langacker 1987, Bybee 1995, etc.). Further, the communicative function of language also plays a crucial role in these accounts. Thus, Tomasello (2003) stresses that children focus on the acquisition of whole utterances and constructions, because utterances are" the primary reality of language from a communicative point of view" (Tomasello 2003: 326).

N. Gagarina and I. Gülzow (eds.), The Acquisition of Verbs and their Grammar: The Effect of Particular Languages, 1–11.

Formal theories, in contrast, tend to schematize a child language in terms of very abstract categories and formal representations, and they start from the premise that a child has (knowledge of) linguistic structures from birth. We suppose that in the present state-of-the-art of language acquisition science one cannot unequivocally and definitively solve the problems of this discussion. It seems that different elements of the *language faculty* (and extralinguistic reality) play more or less prominent roles in the different levels of the acquisitional process and, consequently, diverse explanatory mechanisms accounting for different phases in the acquisition of language, may be of stronger or weaker force.

Despite the different views on the nature of language, both theoretical approaches agree on the central role of verbs in a sentence and in the grammar.[1] The contributions comprising this volume come from different theoretical backgrounds including the two contrasted above approaches and cover a wide range of issues in *verb grammar;* they all address the main issue of the book: the impact of specific languages on the acquisition of verbs and their grammar. Language specificity causes variation from the onset of children's verbal behavior and affects the domains of (verb) grammar during the whole course of language acquisition. Although empirical evidence is given for both language-specific and more general processes (e.g. the developmental stages) in the acquisition of verbs, language-specific performance is more substantially observed and is registered from an age when the rule-governed behavior of children is just starting to develop (cf. for example, Bowerman and Levinson 2001, Naigles and Lehrer 2002). The contributors' original empirical data show that rule-governed behavior and rule-based accounts of the acquisition process become essential later in development, when a certain number of abstract paradigmatic and syntagmatic operations have already been mastered. Within this major topic, which constitutes the 'red-thread' of this volume, several sub-issues define the further organization of the contributions. The first deals with the bootstrapping effect, the second explicates the role of frequency in the acquisitional process, and the third compares learner-specific courses of development within a single language.

The main goal of the book is thus not to confront the different acquisitional theories, but a) to bring more empirical longitudinal data on various languages into the discussion and to show both similar and divergent tendencies in these data, b) to show how different theoretical approaches interpret these data and account for the development of language (use and grammar), c) to demonstrate the validity of both approaches and their potential for compatibility, and in this way move theoreticians forward in their understanding of the phenomenon of language acquisition. The two latter goals seem to be best achieved by investigating an issue which is familiar to both theoretical approaches, namely, the development of

language-specific competence within the acquisitional process. Together with the focus on acquisition of the verb and its grammar research in this domain provides a fruitful basis for discussion.

The maturation model of language acquisition assumes that UG becomes the language specific grammar over time and that UG is entirely available only up until the time when the native language has been completely acquired (cf. Atkinson 1992, Wexler 1999). Constructivist models that may also be opposed to theories of UG alongside with the usage-based approaches mentioned above mostly elaborate on the early acquisition of spatial relations (e.g. Bowerman and Choi 2001, Sinha et al. 1999); however, two main hypotheses of this approach – a holistic view of universal spatial cognition and the language specific acquisition hypothesis – are beyond the main scope of this book.

The book presents original contributions based on analyses of naturalistic data from eleven languages: Croatian, Dutch, English, Estonian, French, German, Hebrew, Jakarta Indonesian, Japanese, Russian and Spanish. Three of the contributions make cross-linguistic comparisons – between English and Russian; English, German and Spanish; and German, Croatian and English. All papers in the volume investigate first language acquisition and one paper studies both first and second language acquisition. Different stages of ontogenesis are presented as well: from the first words, when word classes are not yet formed, to the period when children produce multi-word sentences with subjects, objects, and complements, and when the morphology of verbs can basically be said to have been acquired.

Part 1: Language-specific impact on the acquisition of Hebrew

The volume begins with two articles investigating Hebrew. The first one treats verbal agreement from a generativist perspective and the second one deals with the acquisition of verb argument structure from a constructivist perspective. These two contributions thus represent opposing viewpoints and reflect the ongoing debate on the nature of the acquisitional process: 'grammar-discovery' (e.g. Pinker 1984, 1989; Gleitman 1990) vs. 'grammar-construction' (e.g. Tomasello 1992, 2003, Slobin 1995). Differences in theoretical explanations of the language acquisition process have their roots in the way that a researcher characterizes a target language. However, neither approach denies the impact that language-specific aspects have on the acquisitional process. While **Uziel-Karl** stresses the role of input and language use, **Armon-Lotem** leaves this issue open. These two theoretically different contributions show that the language-specific nature of Hebrew is more strongly reflected in the later phases of the acquisitional process when children move from concrete item-based entities to more general rule-based structures. The authors agree on the fact that, in the beginning, children do not possess grammatical knowledge about their language and either "restrict

their verb use to specific lexical items" (Uziel-Karl) or "generate smaller convergent trees" (Armon-Lotem). Within the course of language development children augment the diversity of argument types, as outlined by the first author, and/or enlarge the set of functional categories within a tree, as argued by the second author. In general, the increasing importance of the language specific structures and rules children have to acquire is presented in both papers, but the sources of growing linguistic competence and the mechanisms underlying this development are treated differently.

Specifically, **Uziel-Karl** argues for an item-based development of verb argument structure in a 'bottom-up fashion' and establishes three developmental phases in L1 acquisition of Hebrew: *Input-based*, *Bottom-up Construction of Generalizations*, and *from Generalizations to Rules*. Verb-specific 'utterance schemas' employing different nouns and increasing diversity of argument types follow the establishment of argument positions. This in turn leads to further development of the variation of verbs and induces further developmental steps on more abstract linguistic levels. Several issues which are relevant for further research, like the transition and overlap between developmental phases (cf. Kilani-Schoch et al. 1997 for the problems of demarcating the early stages of morphological development), the sensitivity of children to certain cues in the input at different developmental phases, etc. are sketched in the conclusion. **Armon-Lotem** analyses early utterances containing verbs and overt subjects and explores the impact of early verbal morphology on the acquisition of dependent verb constituents in Hebrew. She uses the "no Agr nodes" approach to account for the interrelation between the early appearance of verbal morphology and the early appearance of subjects and argues for a central role for verbs in the acquisition of phrase-markers. In particular, she highlights some agreement mismatches resulting from the overlap between a) the order of acquisition of agreement morphology (agreement markers for gender and number < appropriate use of tense morphology < agreement markers for person) and children's early tendency to mark telic verbs for past tense and atelic verbs for present tense and b) the onset of the production of overt subjects.

Part 2: Language-specific variation in the development of predication and verb semantics

Despite the increasingly discussed phenomenon that 'grammatical structures are remarkably similar across languages' (Polinsky 2004: 437) one can find notable variation in the process of L1 acquisition. This variation comes to light and develops together with a child's increasing competence in his native language. The articles in this part give an example of the outstandingly language-specific manner of acquisition within the domain of predication and

of verb semantics. The paper of **Czinglar, Katičić, Köhler and Schaner-Wolles** investigates copular(less) constructions in English, German and Croatian. These three languages differ with respect to their allowances for predication structures with and without a copula. No copula omission has been found for Croatian, whereas the two other languages exhibit a tendency for omission of the copula within certain predicate types, namely, locative (temporally bounded) predicates. The authors argue for interrelatedness between the presence of copulaless predication structures and the *root infinitive* phenomenon. In addition, the overwhelming majority of analyzed sentences with copulas in English and German contain the finite variant. **Tsujimura's** paper shows that the L1 Japanese-speaking child does not learn all verbs with equal ease, but has a bias toward (morphologically marked) intransitive verbs. Early verbs predominantly denote motion with special path or definable goal. A language-specific principle focusing on result, rather than process, in describing an event is taken to be a cornerstone for this lexical bias. The author comes to general conclusions concerning the acquisition of language-specific features and their role in conceptual development.

Part 3: Stages in the development of verb grammar and the role of semantic bootstrapping

The contributions in this part deal with the bootstrapping mechanism in the acquisition of verb semantics and verb inflection and with the evaluation of periods in the acquisition of the verb and its forms. The notion of bootstrapping has been traditionally employed in formalist theories of language acquisition as a 'step-link' assisting children in learning new elements of grammar and in establishing connections between various grammar domains (cf. Weissenborn and Höhle 2001). Originating from Pinker's (1984) *semantic bootstrapping* the concept of *bootstrapping* was extended to prosodic/phonological (Morgan and Demuth 1996), morphological and syntactic areas and is now used as an explanatory device outside generativist models of language acquisition. In this part of the book, the role of a bootstrap for the development of verb semantics and verb grammar is shown within the different periods of language acquisition. **Herr-Israel and McCune** observe the single word period when word classes are in the process of being formed and children use relational words as a bootstrap for verb meanings and early word combinations. The authors argue for a special single word period which they consider to be a crucial phase in the development of the use and understanding of verbs in multiword combinations. As shown in the paper, the investigated children express dynamic actions with relational words within the single word period. Primal

verbs overlap in meaning with relational words and, in the period of combining words, both relational words and primal verbs predominate over other lexical means.

The notion of verb paradigm serves for **Ingram, Welti and Priem** as a basis for their model of verb acquisition. Proposing four stages in the acquisition of verb paradigms in Spanish, German and English, the authors found only partial evidence for the semantic bootstrapping strategy. They suggest that children may differ in the way that they exhibit the transition from one developmental period to another and in the way that they use bootstrapping mechanisms to learn their native language. The article also raises methodological questions in the study of verb acquisition. The authors first propose a feature – verbal syntactic type – to determine a child's verb stage that is not tied to MLU. Secondly, they discuss the methodology of the recording, transcription, and storage of language samples on the basis of new techniques for the digitization of language samples. Their results are especially interesting in light of the fact that the three languages analyzed have different levels of morphological richness. However, all these languages exhibit a similar pattern with respect to developmental paradigms.

Dimroth and Jordens discuss the development of finiteness in L1 and L2 acquisition of Dutch. The understanding of finiteness falls into two parts, concerning the semantic concept of finiteness and its morphosyntactic marking. Given that finiteness carries the pragmatic function of assertion and relates the descriptive content of an utterance to its topic component, the authors propose three successive stages in the acquisition of finiteness. Comparison between L1 and L2 learning revealed similarities consisting of what authors named *validation* (or *linking device*). This expresses the relations between the topic and the state of affairs of an utterance and manifests itself in various ways during the proposed stages of the acquisition of finiteness.

Part 4: Language-specific variation and the role of frequency

Two papers in the fourth part of the volume contribute to the ongoing debate surrounding the role of input frequency in the acquisition process (**Gil, Gülzow and Gagarina**). Recent studies regard input frequency as an influential factor in the order of acquisition and the overall representation of linguistic elements in children's early productions (cf. Rowland et al. 2003, Wijnen et al. 2001). Although no consensus is reached regarding the precise nature of how the input frequency of a certain linguistic element influences the way in which children learn their target language, most studies agree that high frequency contributes to an element's salience (e.g. Bowerman 1990, Snyder and Stromswold 1997, cf. Gropen et al. 1991). In part, the diversity of

opinions expressed may be regarded as a result of the different theoretical perspectives that are adopted; frequency is discussed in input based approaches (for an overview see Rowland et al. 2003, Wijnen et al. 2001) as well as in studies simulating language use (e.g. Elman et al. 1998, Seidenberg and MacDonald 1999), or more generative approaches such as, for instance, Yang (2002), who combines a general learning approach with a UG-based hypothesis of language learning.

In the present volume, some studies follow approaches in which the focus is on the impact of frequency on the overall representation of certain linguistic elements in children's utterances while other studies concentrate on the order of acquisition. Demuth (1989), for instance, has shown that children learning Sesotho as their first language mirror their target language in that they produce a larger number of passive structures than English children do at early stages of language acquisition. **Gil** argues along the same lines that the early and relatively numerous uses of voice prefixes in Jakarta Indonesian is an effect both of the semantic transparency of the prefixes and their high frequency in language addressed to children. The paper by **Gil** explores the acquisition of passive and active voice markers (voice morphology) in longitudinal data from eight children acquiring the Jakarta dialect of Indonesian. It is shown that some children acquire the two voice prefixes *di-* and *N-* very early (before the age of two) and that, generally, the acquisition of passive precedes the acquisition of active voice. The proposed account is based on two distinct factors, one pertaining to the high input frequency of voice prefixes, the other to their morphosemantic transparency.

Similar effects are also noted by researchers studying the acquisition of other linguistic elements: Mesook et al. (1999) demonstrate that L1-English speaking children and L1-Korean speaking children use locative verbs according to the structure of their target language from early on, Gathercole (1986) shows in a study on the acquisition of English that Scottish children use a greater number of verbal constructions involving the past participle than American children, and a comparison of English-speaking and Mandarin-speaking children revealed that the children produce nouns and verbs in context specific but target-like frequencies (Tardif et al. 1999).

Taking a related perspective, **Gülzow and Gagarina** argue for an interaction of the frequency effect with a given linguistic element's semantic transparency or communicative relevance. Comparing the early verb use of Russian and English children, Gülzow and Gagarina find that the Russian and the English children's use of main verb forms mirrors the relevant input frequencies from early on. This results in a more accurate structure for the Russian children's verbal utterances, as they mainly produce the synthetic verb forms of verbal utterances with a single verb form. The English children produce many non-target verbal utterances, using the base form with no other

verbal element. However, all children in the study manifest short periods in which their use of non-finite verb forms is above the ratio of these verb forms in the speech addressed to them. While the degree of overrepresentation varies between the children and the individual verb forms, Gülzow and Gagarina argue that the underlying mechanism is the same: due to the increased salience (e.g. sentential position, prosodic stress) and frequency of non-finite verbal elements they have a greater impact on the process of language acquisition for a given period of time.

In their study comparing the acquisition of Estonian in two boys, **Vihman and Vija** (see part 5 below) show that frequency may be responsible for the early emergence of certain verbal elements while other elements seem to remain unaffected. Estonian *on* 'is' is represented in high numbers in the input of one boy that Vihman and Vija studied and appears relatively early in the boy's productions. The ratio of *ma*-infinitives versus *da*-infinitives, on the other hand, is not mirrored in the early language of the Estonian boy. A central aspect of the frequency debate that Vihman and Vija draw attention to is the attempt to pin down how frequency interacts with other properties that contribute to the salience of a given linguistic element. Type and token frequency alone do not explain why certain verb forms appear early while others do not. In Vihman and Vija's study, which compares a monolingual boy with a bilingual boy acquiring Estonian and English, it is possible to relate the boys' different paths to the parallel learning of two languages by the bilingual child.

Part 5: Language-specific and learner-specific peculiarities in the development of verbs and their grammar

Three contributions of the last part present data analyses of children acquiring Estonian, French and Russian as their L1. The papers deal with the emergence of verbs and their (contrastive) finite forms and find significant learner-specific variation in the course of language acquisition. It is shown that language- and learner-specific factors overlap but do not substantially affect the general path of language acquisition presented in the other chapters of the book. Analyzing the data of two boys acquiring Estonian, monolingual Andreas and bilingual Raivo, **Vihman and Vija** stress the fact that one child focuses on regular forms and productive 'rule-based' patterns while the other child instead focuses on irregular forms and item-based learning. The authors argue that learning only one vs. two languages is not necessarily the reason for such linguistic behaviour since another monolingual child could be shown to follow the path of the bilingual one.

The study of **Martinot** examines the significance of first verbs in the development of grammar in French and thus approaches issues of the interface between the development of verbs and the child's entire grammar. She addresses questions dealing with the prevalence of certain verbal forms and a) their comparative prevalence among the children investigated and b) their semantic conceptualization and the role they play in argument construction. Martinot finds that although children show individual differences in lexical diversity, they use language-specific light verbs, which play a crucial role in the process of language acquisition. Her analysis shows parallelism in morphological and lexical-grammatical development, even though the origin of acquisitional processes remains unclear.

Finally, **Kiebzak-Mandera's** paper studies the emergence of finite verbal forms marked for person in Russian, with special emphasis on their function. The production of personal forms of verbs is compared with the usage of forms which are not marked for person, such as past and non-finite forms. The two children investigated show remarkable differences in coding speaker and addressee. These differences reflect two strategies existing in the target language: to name the addressee with first/second person forms or with the third person form. Rare and late occurrences of second person forms are not due to difficulties in learning but one due to the fact that children seldom address the interlocutor.

The contributions to this volume discussing different issues in the acquisition of verbs and their grammar in various languages provide an overview of their lexical, morphological, and syntactical peculiarities. We hope that this will, implicitly, provide a road map that leads the reader to the blind spots of our knowledge of how children discover the language-specific world of verbs. By presenting different theoretical accounts for a range of naturalistic data we have tried to sketch possible ways towards better understanding of (specific) language acquisition phenomena.

1. NOTES

[1] Gentner and France (1988) posit the verb centrality under question.

2. REFERENCES

Atkinson, Martin. *Children's Syntax: An Introduction to Principles and Parameters Theory*. Oxford: Blackwell, 1992.

Barlow, Michael and Suzanne Kemmer. *Usage-based models of language*. Stanford: CSLI Publications, 2000.

Bowerman, Melissa and Stephen C. Levinson (ed.) *Language Acquisition and Conceptual Development*. Cambridge: Cambridge University Press, 2001.

Bowerman, Melissa and Soonja Choi. Shaping meanings for language: universal and language specific in the acquisition of spatial semantic categories. In Melissa Bowerman and Steven Levinson (eds.), *Language Acquisition and Conceptual Development.* Cambridge: Cambridge University Press, 2001: 475-511.

Bybee, Joan. Regular morphology and the lexicon. *Language and Cognitive Processes* 10 (1995): 425-455.

Demuth, Katherine. Subject, topic and the Sesotho passive. *Journal of Child Language* 17 (1989): 67-84.

Elman, Jeffrey L., Elizabeth A. Bates, Mark H. Johnson, Annette Karmiloff-Smith, Domenico Parisi and Kim Plunkett. *Rethinking Innateness. A Connectionist Perspective on Development*. MIT Press, 1998.

Gathercole, Virginia. The acquisition of the present perfect: explaining difference in the speech of Scottish and American children. *Journal of Child Language* 6 (1986): 423-42.

Gleitman, Lila. The structural source of verb meaning. *Language Acquisition* 1 (1990): 3-55.

Gropen, Jess, Steven Pinker, Michelle Hollander and Richard Goldberg. Syntax and Semantics in the Acquisition of Locative Verbs. *Journal of Child Language* 18 (1991): 115-151.

Kilani-Schoch, Marianne, Anna de Marco, Anastasia Christofidou, Maria Vassilakou, Ralf Vollmann and Wolfgang U. Dressler. "On the demarcation of phases in early morphology acquisition in four languages." In Katarzyna Dziubalska-Kołaczyk (ed.), *Pre- and Protomorphology in Language Acquisition*. Papers and Studies in Contrastive Linguistics 33, Poznań: Adam Mickiewicz University, 1997: 15-32.

Langacker, Ronald W. *Foundations of Cognitive Grammar.* Vol. I. Stanford: Stanford University Press, 1987.

Meesook, Kim, Barbara Landau and Colin Phillips. Cross-Linguistic Differences in Children's Syntax for Locative Verbs. *BUCLD 23: Proceedings of the 23rd annual Boston University Conference on Language Development*, (1999): 337-348.

Morgan, Jim and Katherine Demuth (eds.). *Signal to Syntax: Bootstrapping from Speech to Grammar in Early Acquisition*. Mahwah, N.J.: Lawrence Erlbaum Associates, 1996.

Naigles, Letitia R. and Nadine Lehrer. Language-general and language-specific influences on children's acquisition of argument structure: a comparison of French and English. *Journal of Child Language* 3 (2002): 545-566.

Newmeyer, Frederick J. Grammar is grammar and usage is usage. *Language* 4 (2003): 682-707.

Pinker, Stephen. *Language Learnability and Language Development*. Cambridge, Mass: Harvard University Press, 1984.

Pinker, Steven. *The language instinct: How the mind creates language*. New York: Morrow, 1994.

Polinsky, Maria and Eric Potsdam. Malagasy Control and its Theoretical Implications. *BLS* 30 (2004).

Rowland, Caroline, Julian M. Pine, Elena Lieven and Anna L. Theakston. Determinants of acquisition order in wh-questions: re-evaluating the role of caregiver speech. *Journal of Child Language* 30 (2003): 609-635.

Seidenberg, M.S. & MacDonald, M.C. A probabilistic constraints approach to language acquisition and processing. *Cognitive Science* 23 (1999): 569-588.

Sinha, Chris, Lis A. Thorseng, Mariko Hayashi, and Plunkett, Kim. Spatial language acquisition in Danish, English and Japanese. In Peter Broeder and Jaap Murre (eds.) *Language and Thought in Development*. Tübingen: Gunter Narr, 1999: 95-125.

Slobin, Dan (ed). *The cross-linguistic study of language acquisition*. Vol.I. Mahwah, N.J.: Lawrence Erlbaum Associates, 1985.

Tardif, Twila, Susan A. Gelman, and Fan Xu. Putting the "Noun Bias" in Context: A Comparison of English and Mandarin. *Child Development* 70 (1999): 620-635.

Tomasello, Michael, Ann Cale Kruger and H. H. Ratner. Cultural learning. *Behavioral and Brain Sciences* 16 (1993): 495-552.

Tomasello, Michael. *First verbs: A case study of early grammatical development*. Cambridge: Cambridge University Press, 1992.

Tomasello, Michael. *Constructing a Language.* Harvard: Harvard University Press, 2003.

Weissenborn, Jürgen and Barbara Höhle (eds.). *Approaches to Bootstrapping: Phonological, Lexical, Syntactic and Neurophysiological Aspects of Early Language Acquisition*, Vol. 1 and 2. John Benjamins Publishing Company, 2001.

Wexler, Kenneth. Maturation and Growth of Grammar. In: *Handbook of Child Language Acquisition*. Ed. by William C. Ritchie and Tey K. Bhatia. San Diego: Academic Press, 1999: 55-105.

Wijnen, Frank, Kempen, M. & Gillis, Steven. Root infinitives in Dutch early child language: an effect of input? *Journal of Child Language* 28 (2001): 629–660.

Yang, Charles D. *Knowledge and Learning in Natural Language*. Oxford: Oxford University Press, 2002.

Part 1

Language-specific impact on the acquisition of Hebrew

SIGAL UZIEL-KARL

ACQUISITION OF VERB ARGUMENT STRUCTURE FROM A DEVELOPMENTAL PERSPECTIVE

Evidence from Child Hebrew

Abstract. The present study proposes a developmental account of Verb Argument Acquisition based on analysis of longitudinal data from four Hebrew-speaking children aged 1;5 - 3;0 years. The following developmental trajectory emerges from the data: Verb argument structure is initially acquired in a bottom-up fashion, limited to specific lexical items. Children start out by rote-learning particular unanalyzed verb forms and verb-argument combinations (holophrases). Then, they start hypothesizing about argument positions, as attested by the use of "groping patterns" (Braine 1976). Once argument positions are set, children start producing verb-specific "utterance schemas" with a wide range of nouns. Gradually, they increase the variation in argument types. As children experience more and more verbs in a variety of communicative contexts, acquisition becomes more abstract and top-down. This is evident from their use of innovated verbs in familiar argument patterns, and from their overextension errors. The proposed account is in line with other input-based accounts of verb argument acquisition (e.g., Braine 1976; Tomasello 1992, 2000a,b).

1. INTRODUCTION

The present paper is concerned with how children develop their knowledge of Verb Argument Structure (VAS).

Two main groups of accounts relate to this question: *Inside-Out* and *Outside-In* (Hirsh-Pasek and Golinkoff 1996). *Inside-Out* accounts, represented by work like Pinker's (1984, 1989) "semantic bootstrapping" and Gleitman's (1990) "syntactic bootstrapping" accounts, contend that children are innately endowed with domain specific linguistic knowledge that enables them to properly map linguistic forms and functions. These accounts argue for a strong continuity between child and adult grammars, and emphasize grammar discovery rather than grammar construction.

In marked contrast, *Outside-In* accounts of various orientations assume no *a priori* language structure, and emphasize the role of input, actual language use, and communicative events in determining the child's linguistic skills. These accounts focus on acquisition as an ongoing PROCESS, and contend that language acquisition takes place by means of domain-general procedures, initially in a bottom-up fashion (Bates and MacWhinney 1987,

N. Gagarina and I. Gülzow (eds.), The Acquisition of Verbs and their Grammar: The Effect of Particular Languages, 15–44.

1989; Bowerman 1973, 1982; Bruner 1983; Maratsos and Chalkley 1981; Ninio 1988; Ninio and Snow 1988; Schlesinger 1988; Tomasello 1992, 2000a,b).

Against this background, the present study proposes a developmental account of VAS acquisition based on longitudinal data from child Hebrew (see also Uziel-Karl 2001). The proposed account consists of three developmental phases: An initial **Input-based** phase, an intermediate phase of **Bottom-up Construction of Generalizations**, and a final phase **from Generalizations to Rules**. These are illustrated in Figure 1.

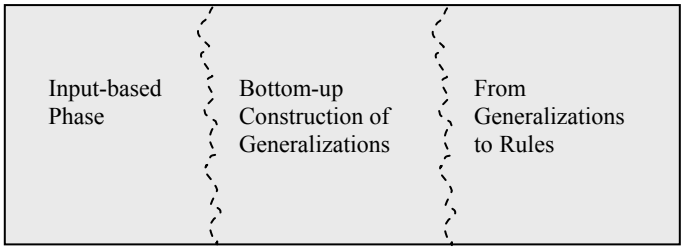


Figure 1. *Developmental Phases in VAS Acquisition*

The first two phases are "verb-bound", as they revolve mainly around the verb – its morphology, semantics, and initial argument structure. They mark the transition from rote-learning to rule formation. The **input-based** period involves mainly accumulation of input to serve as the basis for subsequent generalizations. During this period, verbs and verb-containing expressions are rote-learned as unanalyzed amalgams in that children do not seem to be aware of their internal structure yet (MacWhinney 1978, 1982; Bowerman 1982). For example, children acquiring Hebrew as their first language seem to acquire verbs like ***raxac*** 'wash' (P1 *pa'al*) and *hit**raxec*** 'wash oneself' (P4 *hitpa'el*) as two separate verbs, without realizing that they share the same root ***r-x-c***, conjugated in two different verb-patterns (Berman 1986).

The period of **Bottom-up construction of generalizations** forms a link between the initial and subsequent periods of VAS acquisition. During this period children show evidence for preliminary organization and analysis of linguistic data, but they still do not formulate rules. In this sense, the initial organization of input into structure is a process of approximation, or schema formation (Bybee and Slobin 1982) that proceeds in a bottom-up fashion (Berman 1993a; Braine 1976; Schlesinger 1988; Tomasello 1992).

The third phase, **from generalizations to rules**, relates to the "verb at large", i.e., to the verb and its expanded argument structure. Before this period, children tend to replicate the structures modeled by individual verbs

in their repertoire. After they have accumulated sufficient data, and generalized it, they start to formulate rules. From this point on, acquisition proceeds top-down, as children associate abstract argument structures from their repertoire with new verbs that enter their lexicon. During this period, innovations and overextensions occur, to be resolved as children encounter more exemplars while at the same time becoming more proficient in other relevant linguistic modules like morphology and semantics. In this sense, this period marks a shift from partial to full productivity in VAS acquisition.

The proposed account is in line with other *Outside-In* accounts of VAS in emphasizing the role of input and actual language use in early acquisition (e.g., Braine 1976; Tomasello 1992, 2000a,b). Likewise, it assumes that children initially operate within different psycholinguistic units than adults, and that VAS acquisition is initially verb-dependent, and proceeds in a bottom-up fashion (cf. also Armon-Lotem; Ingram, Welti and Priem, this volume).

2. METHOD

To support my account, I use naturalistic longitudinal speech samples collected on a weekly basis from four Hebrew-speaking children, 3 girls (Hagar, Smadar and Lior), and a boy (Leor), aged 1;5 - 3;0 years. The corpus from which my data were extracted was recorded and transcribed as part of the Crosslinguistic Language Acquisition Project conducted by Berman and Weissenborn (1991).

The children were audio-recorded at home during interactions with their parents and siblings. Each child was recorded for approximately one hour a week in a variety of situations, and in more than one session. The girls were recorded by their mothers, and the boy by his aunt, all graduate students of Linguistics at Tel Aviv University at the time.

The present study analyzes data from transcripts of sessions recorded twice a month, at intervals of 10-14 days. These intervals are sufficiently short not to miss significant developmental changes in the children's language, yet extended enough to allow such changes to take place. Information about the subjects and database is summarized in Table 1.

Table 1. Subjects and Database

Child	*Sex*	*Age Range*	*No. of Transcripts*	*Mean Utterances per Transcript*
Lior	Female	1;5-3;0	32	168
Smadar	Female	1;5-2;4	13	230
Hagar	Female	1;7-3;0	35	173
Leor	Male	1;9-3;0	32	203

Table 1 shows that the mean number of transcripts per child is close to 30 (M = 28.5), and the mean number of utterances per transcript is close to 200 (M = 193.5), which make the database large enough to draw generalizations.

The data were transcribed and coded using CHILDES (MacWhinney 1995) with adaptations to Hebrew. Analysis included all verb-containing utterances in the children's data, with the exclusion of unintelligible utterances, and utterances which did not contain a lexical verb.

All verbs in the database were coded for lexical category (verb), lexeme (root + verb-pattern), and inflectional morphology (number, gender, person and tense/mood), as shown in example (1).[1]

(1)

Child utterance:	*boi*
Gloss:	come-2S:FEM:IMP
Lexical coding:	V:*bwa1*
Morphological coding:	V:2S:FEM:IMP

In example (1), the verb *boi* 'come-2S:FEM:IMP' is lexically coded as V:*bwa*, where *V* stands for *Verb*, *bwa* is the verb's root, and *1* stands for the *pa'al* (P1) verb-pattern (see section 3.3.2). Morphologically, the verb is in the 2nd person, singular, feminine, imperative form.

All utterances that contained a lexical verb were coded for potential and realized argument structure. *Potential argument structure* refers to idealized, fully spelled-out sets of argument structures that include all obligatory arguments required by a particular verb. For example, the potential argument structures for the verbs *give*, *wash*, and *arrive* are SVOI, SVO and SV, respectively. Coding of potential argument structure was based on prior analyses of VAS in Hebrew (Berman 1982; Armon-Lotem 1997; Stern 1979, 1981), and on my intuitions as a native speaker.

The potential argument structure of a verb may be realized in full or in part in different conversational contexts. Thus, coding of realized argument

structure was based on children's actual production data. A sample coding is provided in (2).

(2) Child Utterance: *roce mayim*
Gloss: want-S:MASC:PRES water
VAS: SVO:EVO

In example (2), the potential argument structure of the verb *roce* 'want' is SVO, and the realized argument structure is VO. The *E* stands for Empty, since the child used the verb without an overt subject.

Each verb-containing utterance was also coded for whether it was self-initiated by the child, exact imitation, or partial imitation of parental input. Children's utterances were classified as partial imitations if they varied from adult utterances in verb morphology (3), or in number of arguments (4).

(3) Adult: *azarti lax*
help-**1S:PAST** to you-2S:FEM
'I helped you'

Child: *la'azor lax*
help-INF to you-2S:FEM
'to help you (= me)'

(4) Adult: *ata roce le'exol?*
you-2S:MASC want-S:MASC:PRES eat-INF
'Do you want to eat?'

Child: *le'exol*
eat-INF
'to eat'

In example (3), the adult uses the verb 'help' in the past tense, while the child uses it in the infinitival form. In example (4) the adult uses an overt subject and a complex verb 'want to eat', while the child reproduces only the main verb 'to eat', and leaves out the subject and modal verb 'want'.

Finally, the data were coded for errors relating to VAS such as word order and overextension errors (for the latter see examples (23)-(27)).

Productivity was measured by: (1) the consistent occurrence of self-initiated, correctly ordered verb-argument combinations for a particular verb across at least three consecutive sessions; (2) use of different arguments (i.e., type variation) with a particular verb across consecutive sessions; or

(3) occurrence of a particular argument structure with two or more verbs across at least three consecutive sessions.

3. RESULTS

This section presents evidence from child Hebrew to support the developmental trajectory outlined at the outset of this study. The section is divided into three parts: An Input-based phase (3.1), Bottom-up Construction of Generalizations (3.2), and from Generalizations to Rules (3.3).

3.1. An Input-Based Phase

Several kinds of evidence support the initial Input-based period of VAS acquisition as characterized above. These include: the distribution of early verb-forms (3.1.1), the early use of "holophrases" (3.1.2), and the effects of parental input on early acquisition of verbs and VAS (3.1.3).

3.1.1. Distribution of Early Verb Forms

Researchers in the field of language acquisition argue that young children's grammatical knowledge is initially organized around specific lexical items (Akhtar 1999; Akhtar and Tomasello 1997; Clark 1995; Lieven, Pine and Baldwin, 1997; Tomasello and Brooks 1999). Similarly, the following data from child Hebrew suggest that children initially rote-learn certain verbs as unanalyzed amalgams in that each verb is acquired in a particular morphological form (i.e., unique number, gender, person, and tense configuration), and children are not aware of its internal structure (i.e., root + verb-pattern).

Table 2 shows the distribution of preferred verb forms (i.e., particular number, gender, person and tense configuration) for six verbs that were among the earliest and most frequently used by the children in my sample between ages 1;5 - 1;11, when they were in a transition from the single-word stage to multiple word combinations (see also Berman and Armon-Lotem 1996). The Table specifies the total number of tokens for each verb, and the morphological form in which the four children (combined) used it most frequently.

Table 2. Distribution of Verb Forms for Six Early Verbs in the Children's Data [1;5 - 1;11]

Verb Form	*Verb Morphology*	*No. of Tokens*	*Distribution of Preferred Form*
tni 'give'	2S:FEM:IMP	20	70%
nafal 'fall'	3S:MASC:PAST	43	93%
bo/*boi* 'come'	2S:MASC/FEM:IMP	62	68%
roce/roca 'want'	S:MASC/FEM:PRES	209	99%
gamarnu 'finish'	1P:PAST	35	77%
(la)sim 'put'	INF/2S:MASC:IMP	64	89%

Table 2 reveals that until around age 1;11 when there is evidence for initial productivity in use of grammatical subjects and morphological inflections, each verb is used in a particular morphological form. For example, Lior initially uses the verb *bwa1* 'come' as *bo* in the imperative masculine form even when referring to her mother. She does not alternate the gender of the verb by the context of use. She uses the verb *npl1* 'fall down' as *nafal* in the 3rd person masculine singular to refer to everything that falls down be it feminine, masculine, plural or singular. She uses the verb *ntn1* as *tni li* 'gimme' in the imperative feminine form, and the verb *rcy1* 'want' as *roca* in the feminine singular present tense. She uses the verb *gmr1* 'finish, end' as *gamarnu*, in the 1st person plural past tense, and the verb *ily1* 'go up' as *la'a lot* in the infinitive in all contexts. Smadar uses the forms *shev* 'sit down' and *sim* 'put' repeatedly to refer to her mother (e.g., *shev ima* 'sit down mommy', *ima sim (mi)ta sus* 'mommy put bed horse = mommy put the horse on the bed') although these forms, if analyzable at all, are closest to the masculine singular imperative form (cf. *sim* 'put-2S:MASC:IMP', *shev* 'sit-2S:MASC:IMP'). Thus, the vast majority of occurrences of each verb are in a single preferred morphological form with almost no alternations or governing rules, regardless of the agreement and tense marking required by the context (see also Berman and Armon-Lotem 1996; Uziel-Karl 1997). The lack of variation in verb morphology throughout the examined period suggests that these verbs are initially rote-learned as unanalyzed elements. These findings corroborate research on the early acquisition of verbs in other languages like Italian (Pizzuto and Caselli 1992, 1993), Spanish (Gathercole, Sebastian and Soto 1995), and Inuktitut (Allen 1996).

The data also suggest that there is no correlation between a verb's initial morphological form and its transitivity value or semantic class. That is, not all transitive or all intransitive verbs are used in the same inflectional configuration. For example, the verbs *rcy1* 'want' and *gmr1* 'finish' which are both transitive, are used with different tenses. Similarly, verbs which

share a semantic class are not necessarily acquired with the same morphological form, e.g., the verbs *ntn1* 'give' and *sym1* 'put' which are both transfer verbs, are used with different tenses. These findings suggest that Hebrew-speaking children initially acquire each of these verbs individually.

Two exceptions need to be accounted for – the use the verbs *rcy1* 'want' and *sym1* 'put'. The verb *rcy1* 'want' is used both in the masculine (*roce*) and feminine (*roca*). Leor, the boy, uses only the masculine form a large number of times, while the girls Hagar, Smadar and Lior use only the feminine. All uses of the verb are in the present tense. The gender of *rcy1* 'want' in the children's data has to do with their sex, as follows. At its early phases, acquisition is assumed to depend to a large extent on parental input. In Hebrew this input differs according to the sex of the addressee, so that boys are addressed in the masculine and girls in the feminine. As a result, boys produce the verb in the masculine form first, while girls initially produce it in the feminine.[2] In the case of *sim/lasim* 'put' these two forms can be attributed to a certain degree of ambiguity, since *sim* could be either a bare infinitive, without the infinitival marker *le-* 'to' or the masculine singular imperative. Since the period of early verbs is transitory with respect to the use of unclear forms, some occurrences of *sim* could be taken to be truncated versions of *lasim* 'to put'.

The following two Tables compare the number of different verb forms for each child across two age periods 1;5 - 1;11 (Table 3) and 2;0 - 3;0 (Table 4). In the present analysis, for any given verb, 2S:**MASC**:IMP and 2S:**FEM**:IMP and 1S:US:**PAST** and 1S:US:**FUT** constitute four distinct verb-forms. In contrast, repeated occurrences of a particular form like S:**MASC**:PRES are counted as a single form.

Table 3. Distribution of Verb Forms for Four Children [1;5 - 1;11]

Lexeme	*Smadar*	*Lior*	*Leor*	*Hagar*
ntn1 'give'	1	2	—	1
npl1 'fall'	3	3	—	2
bwa1 'come'	4	3	4	4
rcy1 'want'	2	1	4	1
gmr1 'finish'	—	1	—	2
sym1 'put'	2	—	3	2

Table 3 reveals that the overall number of distinct verb forms for each of the six verbs (*ntn1* 'give', *npl1* 'fall', *bwa1* 'come', *rcy1* 'want', *gmr1* 'finish', *sym1* 'put') is initially very small. Yet, certain verbs, e.g., *bwa1* 'come' tend to occur in different verb forms earlier than others, e.g., *gmr1*

'finish'. This may be attributed to the semantic properties of the verb (e.g., semantic class, specificity) rather than to its morphological (e.g., verb-pattern) or syntactic properties (e.g., verb argument structure). For example, the verb *bwa1* 'come' belongs to the class of general-purpose verbs, which has been noted to have particular properties that make it favored for early acquisition (Clark 1993; Ninio 1999; Pinker 1989; Uziel-Karl 2000). As a consequence of its early and frequent use, this verb may initially be prone to greater variation than verbs belonging to other semantic classes.

Table 4. Distribution of Verb Forms for Four Children [2;0 - 3;0]

Lexeme	*Smadar*	*Lior*	*Leor*	*Hagar*
ntn1 'give'	10	6	6	10
npl1 'fall'	4	5	11	8
bwa1 'come'	4	8	10	11
rcy1 'want'	4	6	4	7
gmr1 'finish'	4	4	8	6
sym1 'put'	10	7	14	11

A comparison between Tables 3 and 4 reveals that the number of different verb forms for each lexeme increases dramatically with age, and that this trend characterizes all four children. That is, as acquisition proceeds, children stop using verbs as unanalyzed amalgams; rather, a variety of morphological forms are acquired, and verbs occur in different tenses and with different inflectional markers of agreement (see section 3.2.4 for a detailed discussion). This developmental trend appears to be shared across languages, as evident from studies on the acquisition of verb-paradigms in languages like Spanish, English and German (see Ingram *et al.*, this volume).

3.1.2. Early Use of Holophrases

Along with or shortly after children start using bare verbs in particular morphological forms, they start producing "holophrases". Holophrases are frozen phrases that result when children attempt to reproduce entire adult utterances, but do so with only partial success (Barrett 1982). Table 5 shows the distribution of frequently used holophrases in my data before age 2;0.

Table 5. Distribution of Frequently Used Holophrases before Age 2;0

Child	*Holophrase*	*Gloss*	*Use of particular construction*	*Total occurrences of Verb*
Lior	*azor lax*	'help you = me'	69%	29
	loh roca	'don't want'	63%	16
Smadar	*sim po*	'put here'	68%	25
	loh yoda'at	'don't know'	88%	18
Leor	*sagarnu or*	'(we) turned off light'	70%	12
	roce musika	'wants music'	90%	10

Table 5 reveals that each child produces his or her own holophrases. Each holophrase consists of a [verb + particular lexical item] (Noun, Pronoun, Adverb, Negation), which does not change across repeated uses, suggesting that these expressions are used nonproductively. For example, Leor uses the noun *musika* 'music' in 9 out of 10 occurrences of the verb *roce* 'want' (90%) before age 2, and Smadar uses the negation word *loh* 'not' in 16 out of 18 occurrences of the verb *yoda'at* 'know' (88%) during that period.

Additional examples of children's holophrases are listed in (5)-(10) below.

(5) Mother: *ma ze, Lior, ma at osa?*
what this, Lior, what you-2S:FEM do-S:FEM:PRES
'What's this, Lior, what are you doing?'

Lior: *tusa [: at osa]* [3]
you-2S:FEM+do-S:FEM:PRES
'you+do' [1;6]

(6) Mother: *ani e'ezor lax?*
I help-1S:FUT you-2S:FEM
'I'll help you'

Lior: *azor [: la'azor] lax*
help-INF you-2S:FEM
'to help you (instead of me)' [1;7]

(7) Mother: *ma nafal?*
what fall-down-3S:MASC:PAST
'What fell down?'

Lior: *fal [: nafal] la*
fall-down-3S:MASC:PAST to-her
'fell down to her' [1;8]

(8) Lior: *tora [: at roa]*
you-2S:FEM+see-S:FEM:PRES
'you+see' [1;9]

Mother: *ani loh roa, ani loh yoda'at le-ma at mitkavenet, at omeret li: at roa*
I not see-S:FEM:PRES, I not know-S:FEM:PRES
to what you mean-S:FEM:PRES, you
say-S:FEM:PRES to-me: you see-S:FEM:PRES
'I don't see, I don't know what you mean, you say to me: you see'

(9) Hagar: *ni li, ni li [: tni li]*
give-2S:FEM:IMP to-me give-2S:FEM:IMP to-me
'gimme, gimme' [1;9]

(10) Hagar: *bo elay*
come-2S:MASC:IMP to-me
'come to-me' [1;9]

Examples (5) and (8) show that children pronounce some of these early combinations as morpho-phonological amalgams, for example, *torá* 'you + see' instead of *àt roá* 'you see'. Example (6) shows that children do not inflect pronouns for the correct person, as in *azor lax* 'help you'. In this example, Lior repeats the 2nd person pronoun *lax* 'to you-2S:FEM', used by her mother, to talk about herself, instead of using the required 1st person pronoun *li* 'to me' (cf. *azor li* 'help me'). Example (10) shows that children use excerpts from nursery rhymes, e.g., *bo elay* 'come to-me' is part of a nursery rhyme in which a child asks a butterfly to come and sit on her hand. Each holophrase is used with a single verb in a unique morphological form and with a single pronoun. Its constituent order is fixed, and it does not extend to other verbs or to other lexical items. It thus appears that children acquire each holophrase as an unanalyzed amalgam. This lends further support to the claim that early acquisition of VAS is nonproductive, initially based on rote-learning. These data corroborate findings on the acquisition of inflectional morphology, and early word-combinations in other languages (Bowerman 1982; MacWhinney 1978, 1982; Tomasello 2000a,b).

3.1.3. The Effects of Parental Input on Early Acquisition of Verbs and VAS

How can the choice of a particular verb form or [verb + complement] combination be accounted for? One possible explanation involves frequency of occurrence in input to the child. On this account, the child will prefer a particular verb form or a particular [verb + complement] combination if they are most often heard in the input. To test this hypothesis, I examined the distribution of the six early verbs listed in Table 2 (*ntn1* 'give', *npl1* 'fall', *bwa1* 'come', *rcy1* 'want', *gmr1* 'finish', *sym1* 'put') in the children's data and in input to the children before age 2. For expository purposes, the present section presents data only from Lior and Smadar. These girls were chosen, since recording of their data began much earlier than that of Leor and Hagar's, hence more relevant to the claims made here concerning early acquisition.

Table 6 presents the distribution of verb forms for the verb *gmr1* 'finish' in input to Lior and in the girl's production data before age 2;0. The verb *gmr1* was chosen, since it occurred in Lior and her caretaker's data a large number of times (Total number of tokens in caretaker's data N = 39, in Lior's data N = 20).[4]

Table 6. Distribution of Verb Forms for gmr1 'finish' in Lior's Data and in Input to Lior [1;5-1;11]

Verb Form	*Gloss*	*Input (Caretaker)*	*Production (Lior)*
gamarnu	'alldone'	67% (26)	65% (13)
gamart	you finish-2S:FEM:PAST	13% (5)	5% (1)
gamarti	'I'm done'	13% (5)	20% (4)
gamarta	you finish-2S:MASC:PAST	3% (1)	10% (2)
nigmeret	finish-S:FEM:PRES	3% (1)	0% (0)
nigmor	'we'll finish'	3% (1)	0% (0)

Table 6 shows that the tightest fit occurred between the caretaker's use of *gamarnu* 'alldone' and Lior's use of this verb form (shaded in gray). Another match occurred in the use of *gamarti* 'I'm done'. Normally, we would expect a match between caregiver 2nd person verbs in addressing the child, and child 1st person forms in response to the caregiver's queries. In view of that, a match in use of 1st person forms suggests that the child does not engage in adultlike question-answer interaction, but rather imitates the use of a particular verb-form in the input.

Table 7 below shows the distribution of verb forms for the verb *rcy1* 'want' in input to Smadar and in the girl's production data before age 2. The

verb *rcy1* was chosen since it occurred in Smadar and her caretaker's data relatively frequently (Total number of tokens in caretaker's data N= 65, in Smadar's data N = 11).

Table 7. Distribution of Verb Forms for rcy1 'want' in Smadar's Data and in Input to Smadar [1;6-1;11]

Verb Form	*Gloss*	*Input (Caretaker)*	*Production (Smadar)*
roca	want-S:FEM:PRES	90% (59)	73% (8)
roce	want-S:MASC:PRES	8% (5)	27% (3)
rocot	want-P:FEM:PRES	2% (1)	0% (0)

In Table 7 the tightest fit occurred between the caretaker's use of *roca* 'want-S:FEM:PRES' and Smadar's use of this verb form (shaded in gray). Similar findings emerged for other verbs by all four children and their respective caregivers. For example, over 60% of all occurrences of the verb 'come' in the children's data between ages 1;5 - 1;11, and in their respective caregivers' input during that period were in the imperative (*bo/boi* 'come-2S:MASC/FEM:IMP'). Similarly, the bulk of occurrences of 'fall' were in the past (*nafal* 'fall-3S:MASC:PAST'), and of 'make/do' in the present (*ose/osa* 'make/do-S:MASC/FEM:PRES'). Matches in verb forms were not characteristic only of verbs that all children and caretakers shared. For example, Smadar and her mother used the verb 'know' in the present tense (*yoda'at* 'know-S:FEM:PRES'), and the verb 'take' in the imperative (*kxi* 'take-2S:FEM:IMP') in over 60% of their occurrences in their data. Similarly, Lior and her mother used the verb 'cry' in the present (*boxe* 'cry-S:MASC:PRES'), and the verb 'help' in the infinitive (*la'azor* 'help-INF') over 60% of the time.

These findings should be considered with the following reservations. *Input* here is limited to caretaker data recorded in the transcripts in the assumption that it represents the overall input to which the child was exposed during the early phases of acquisition. Also, the number of occurrences of any particular verb at these early phases is relatively small. Nonetheless, the data suggest that there is a similarity between the distribution of particular morphological forms in input to the two girls and their production data.

Consider next the distribution of early [verb + argument] combinations for the ditransitive verbs *ntn1* 'give' and *bwa5* 'bring' in input to Lior and in the girl's production data between ages 1;5 - 1;11 (Table 8). Note, that the vast majority of occurrences of these verbs in the two databases was in the imperative, and as a result included very few cases of overt subjects.

Table 8. Distribution of Early Arguments for bwa5 'bring' and ntn1 'give' in Input to Lior and in the Girl's Production Data [1;5 - 1;11][5]

			Child			*Adult*	
Verb	*Total No. of Tokens*	*VO*	*VI*	*Other*	*VO*	*VI*	*Other*
bwa5 'bring'	A40, C4	40%	20%	40%	65%	23%	12%
ntn1 'give'	A14, C8	25%	75%		14%	86%	

Two major findings emerge from Table 8: First, even though both verbs have the same potential argument structure (Subject, Direct Object, Indirect Object), each one was acquired with a different first argument: *bring* was initially acquired with an overt Direct Object, e.g., *tavii sefer* 'bring-2S:FEM:IMP (a) book', while *give* was initially acquired with an Indirect Object, e.g., *tni li* 'gimme'. Secondly, a comparison of the child and caretaker columns reveals that the initial argument structures of *bring* and *give* in Lior's data matches parental input for the two verbs. This suggests, in turn, that Lior is attentive to her caretaker's input, and that she processes this input to produce similar patterns.

3.2. Bottom-up Construction of Generalizations

The period of *Bottom-up Construction of Generalizations* links the initial *Input-based* period with the period *from Generalizations to Rules*. As a transitional period, its borders are fuzzy. During this period, children start forming generalizations about possible argument positions in their language, but these generalizations seem to be of limited scope, and to be formed bottom-up rather than top-down. This is supported by two main types evidence: (1) Children's use of unanalyzed verb forms with a specific complement interchangeably in pre- and post-verbal positions, akin to Braine's (1976) "groping patterns"; and (2) their use of particular verbs, each with a single open argument slot as in Braine's "positional productive patterns", or Tomasello's (2000a,b) "Utterance schemas". The following sections discuss each type of evidence in turn.

3.2.1. Groping Patterns

Children's preliminary attempts at forming some kind of generalizations occur when they use particular unanalyzed verb-forms with a specific complement interchangeably in pre- and post-verbal position, akin to Braine's (1976) "groping patterns". According to Braine, these alternations occur since "the child is groping to express a meaning before he has acquired

a sufficient set of rules for its expression" (p. 10). Braine notes that children produce these patterns with an apparently free word order, in a small number of combinations, and often with uncertainty and effort. Over time, these patterns are replaced by "positional productive patterns", characterized by non-free word order and productivity. The following pairs of utterances list examples of "groping patterns" from the four children in the sample.[6]

(11) Lior: *ima gamarta*
Mommy finish-2S:MASC:PAST
'Mommy, you finished' [1;11]

gamarta ima
finish-2S:MASC:PAST Mommy
'you finished, Mommy'

(12) Smadar: *Gaga halax*
Gaga go-3S:MASC:PAST
'Gaga went away' [1;6]

halax aba
go-3S:MASC:PAST Daddy
'Daddy went away'

(13) Leor: *yexabe ner*
blow-out-3S:MASC:FUT candle
'(the) candle will blow out' [2;1]

ima ner nexba
Mommy candle blow-out-3S:MASC:PAST
'Mommy, (the) candle blew out'

(14) Leor: *kadur lesaxek*
ball play-INF
'to play ball' [1;11]

nesaxek im kadur
play-1P:FUT with ball
'we'll play with (the) ball'

(15) Leor: *mayim mevia*
water bring-S:FEM:PRES
'bringing water' [2;1]

hem hevi acic
they bring-3S:MASC:PAST flowerpot'
'they brought (a) flowerpot'

(16) Hagar: *Yael hevia le-Hagari sorya ba-dli*[7]
Yael bring-3S:FEM:PAST to Hagari sorya in the bucket
'Yael brought to Hagari sorya in the bucket'

Hagar: *Sorya hevi*
Sorya bring-3S:MASC:PAST
'brought Sorya' [2;2]

(17) Smadar: *nafal domino*
fall-3S:MASC:PAST domino
'dropped (the) domino = (the) domino dropped'

domino nafal
domino fall-3S:MASC:PAST
'(the) domino dropped' [1; 7]

Example (11) shows that Lior uses a bare verb with a nonargument vocative in pre- and post-verbal position. Similar examples occur in Smadar's data for the verb 'take'. Example (12) shows that Smadar uses the intransitive verb 'go' with a pre- and post-verbal subject, despite violation of the canonical (i.e., SVO) word order in Hebrew in the latter case. The verbs *isy1* 'make/do' and *pcc3* 'pop up' were also used in a similar way. Example (13) shows that Leor uses the verb 'blow out' with a pre- and post-verbal subject. Examples (14)-(16) show that Leor and Hagar use the verbs 'play' and 'bring' in two different morphological forms (nonproductive, see lack of subject-verb agreement in example [15] above) with pre-and post-verbal direct- and indirect-object (i.e., in the case of 'with a ball'). Finally, example (17) shows that Smadar uses the unaccusative verb *npl1* 'fall' with a pre- and post-verbal subject as permitted in Hebrew for this type of verbs. And indeed, the latter type of alternations persists after children have fixed the position of arguments for all other verbs in the language. These examples, thus, indicate that children start hypothesizing about the possible positions of verb-complements in their language.

Thus, Hebrew-speaking children's "groping patterns" are characterized by the lack of variation in verb-form, and at the same time, by flexibility in verb-complement position. But, do these structural variations (e.g., *Gaga halax* 'went away Gaga', *halax aba* 'Daddy went away') also represent differences in meaning? Careful examination of the conversational contexts in which these variations are used, reveals that, initially, children do not assign a distinct meaning to each structure. So, for example, both 'went away Gaga' and 'Daddy went away' are used by Smadar [1;6] to describe an event of disappearing, and both 'ball play' and 'play ball' are used by Leor [1;11] to express a wish to play ball.

Examination of later structural variations reveals, however, that children do begin to differentiate the various structural configurations of particular verb-complement clusters. They do so by assigning each structural form a different pragmatic function. For instance, in example (16), Hagar fronts the direct-object to pre-verbal position in order to emphasize it. This growing ability to establish a one-to-one relation between syntactic form and pragmatic function corroborates findings reported by Ferdinand (1997) for the acquisition of subjects in French.[8] According to Ferdinand, French-speaking children use dislocations (e.g., *Jean, il est parti/Il est parti, Jean* 'Jean, he left') to mark topic-comment relations, since in L1 acquisition, the development of syntax sometimes precedes the development of pragmatics.

3.2.2. Utterance Schemas

Along with a brief use of "groping patterns", or soon afterwards, children start showing evidence for setting argument positions. This evidence occurs in the form of fixed [verb + one-argument] combinations, initially restricted to particular verbs, differing across individual children. These combinations, or "utterance schemas" seem to result from children's partial success in reproducing adult utterances addressed to them (Tomasello 2000a,b). (18) and (19) below list examples of "utterance schemas" for the verbs *sgr1* 'turn off' and *rcy1* 'want' produced by Leor [1;9]. Similar constructions were found for the other three children. In the examples, the verb is marked in bold to distinguish it from the varying lexical items in argument position.

(18) *sgr1* 'turn off' [V N]

sagarnu *or*
turn-off-1P:PAST light
'(we) turned off (the) light'

sagarnu *sefer*
close-1P:PAST book
'(we) closed (the) book'

sagarnu *ha-meavrer*
turn-off-1P:PAST the fan
'(we) turned off (the) fan'

(19) *rcy1* 'want' [V N]

roce *mayim*
want-S:MASC:PRES water
'wants water'

roce *psanter*
want-S:MASC:PRES piano
'wants (a) piano'

roce *tmuna*
want-S:MASC:PRES picture
'wants (a) picture'

roce *sefer*
want-S:MASC:PRES book
'wants (a) book'

roce *tushim*
want-S:MASC:PRES coloring pens
'wants coloring pens'

In examples (18) and (19), Leor uses each verb in a particular morphological form: *sgr1* 'close, turn off' is used in the 1st person plural past, and *rcy1* 'want' in the singular masculine present form. Each verb is used with a single argument in direct object position, filled in by a large variety of nouns. This is in line with a pattern reported by Braine (1976) for another Israeli girl named Odi, recorded in weekly play sessions from 1;11 to 2;2 months, MLU about 1.4. Odi used the verbs *ntn1* 'give,' and *ray1* 'see' in a particular morphological form, with a single argument. Braine notes that *ten/tni li X* 'give-2S: MASC/FEM:IMP' was used with nouns like *kova* 'hat', *mayim* 'water', *oto* 'car', *ze* 'it', *kacefet* 'whip cream', and *te* 'tea' as a formula for request forms; *tire/tiri X* 'see-2S:MASC/FEM:FUTIMP' was used with *kos* 'glass', *susim* 'horses', *ofanayim* 'bicycle', *rakevet* 'train', *kise*

'chair', *buba* 'doll', and *kova* 'hat' to indicate or identify things. Odi also used *eyn* 'there isn't' (*tipot-af* 'nose-drops', *masmer* 'nail'), and *ose* 'make/do-S:MASC:PRES' (*nadned* 'swing', *bayit* 'house', *brr*) in a few [verb + direct object] combinations.

As the data suggest, these constructions have more grammatical structure than holophrases, since, unlike holophrases, they have at least one open slot that may be filled by many different lexical items. It is this semi-productivity, or "filling in" of their previous holophrases, that serves as evidence for children's early attempts at forming generalizations.

3.2.3. Type Variation

However, to become syntactically competent language users, it does not suffice to set argument positions or to "fill in" argument slots. Children should also be able to "break down" their holophrases (Tomasello 2000a). That is, they need to be able to use different types of arguments with each verb in their lexicon. Example (20) illustrates the gradual expansion of argument types with developmental data from Lior.

(20) *loh roca*
not want-S:FEM:PRES
'don't wanna' [1;9]

ani roca
I want-S:FEM:PRES
'I want' [1;9]

roca sukarya
want-S:FEM:PRES a candy
'(I) want a candy' [2;1]

ani roce xalav
I want-S:FEM:PRES milk
'I want milk' [2;1]

roca lilbosh grabayim
want-S:FEM:PRES to wear socks
'(I) want to wear socks' [2;3]

roca she-texamemi lanu et ha-beyca
want-S:FEM:PRES that you warm-up-2S:FEM:FUTIMP to us ACC the egg
'(I) want you to warm up the egg for us' [2;7]

Example (20) shows that Lior starts off with the holophrase *loh roca* 'don't wanna', which she uses repeatedly as an unanalyzed amalgam (see Table 5). Next, she uses the verb with a single noun complement interchangeably in pre- and post-verbal position. This argument structure is then expanded to include two noun complements both in pre- and post-verbal position. Finally, the number of complement types in post-verbal position is increased to include, apart from noun complements, an infinitival complement and a subordinate clause.

3.2.4 Morphological Development

Soon after children start to "break down" holophrases and to show evidence of expanding their inventory of argument types, they begin to exhibit qualitative and quantitative changes in the development of inflectional morphology.

As noted earlier (section 3.1.1), before these developmental changes occur, the acquisition of inflection consists mainly in children's rote-learning of a small number of individual verbs, each in one inflectional form (e.g., *tni* 'give-2S:FEM:IMP', *nafal* 'fall-3S:MASC:PAST', *gamarnu* 'finish-1P:PAST', *roce/roca* 'want-S:MASC/FEM:PRES'). From this point on, rote-learning gradually gives way to a more productive use of inflection, and eventually, to the application of morphological rules. These developmental changes are initially manifested through the emergence of inflectional oppositions (Kilani-Schoch and Dressler 2000). That is, around age 2;0 children start using verbs in two or more forms that differ from one another in number, gender, person, or tense. For example, at age 1;11 Lior uses two distinct forms of the verb *isy1* 'make/do', *osa* 'make-S:FEM:**PRES**' and *la'asot* 'make-**INF**', each with a different tense inflection. Similarly, at age 1;10 Smadar uses two forms of the verb *yshb1* 'sit', *yoshev* sit-**S**:MASC:PRES and *yoshvim* sit-**P**:MASC:PRES, each with a different number inflection.[9] This change is accompanied by a steady increase in the number of verb-containing utterances in children's production data, as well as in the number of verb-types in their lexicons. So, for example, Lior and Smadar show a dramatic increase in the number of verb types around age 2;0 (Lior) and 1;10 (Smadar), and at about the same age, increase the number of inflectional oppositions per verb (Lior 2;1, Smadar 1;10).

Over time, children go beyond the two-form oppositions, and start using each verb with a variety of inflectional forms, thus forming mini-paradigms (Kilani-Schoch and Dressler 2000). The first true (i.e., self-initiated) mini-paradgim in Lior's data occurs at age 2;3, with the verb *shty1* 'drink', and

consists of four different inflectional forms, and in Smadar's data at age 1;10 with the verbs *isy1* 'make/do' and *akl1* 'eat', each with five inflectional forms (for a detailed description of this process see Uziel-Karl 2004).

The acquisition of inflectional morphology has been noted to proceed along similar developmental steps in different languages (Bittner, Dressler and Kilani-Schoch 2000), and to play a significant role in the acquisition of verbs and verb argument structure cross-linguistically (Ingram *et al*., this volume). This process plays a particularly crucial role in the acquisition of argument structure in a mixed *pro-drop* language like Hebrew, as it also affects the acquisition of argument ellipsis (Uziel-Karl and Berman 2000). In Hebrew, subject ellipsis is licensed only in contexts in which verb inflection is sufficiently rich to ensure that the reference of the missing argument can be reconstructed (Ariel 1991). Thus, children's inflectional system has to be sufficiently rich before they can license argument ellipsis in syntactically appropriate contexts.

3.3. From Generalizations to Rules

With exposure to larger masses of input, children become more consistent in using particular argument structure patterns, and expand these patterns to additional verbs (Akhtar 1999; Akhtar and Tomasello 1997; Clark 1995; Lieven, Pine and Baldwin 1997; Tomasello and Brooks 1999). As more verbs interact with more patterns to achieve a "critical mass" (Marchman and Bates 1994), children abstract away from particular schemas, and their knowledge of VAS becomes increasingly top-down and constructionist rather than bottom-up and lexical. It is during this period that their behavior can first be characterized as rule-bound.

3.3.1. Innovated Verbs in Familiar Argument Structure Patterns

Evidence for children's rule-governed behavior is found in their attempts to associate argument structures from their already established repertoire with innovated verbs they produce. Examples of such innovated verbs from Smadar and Leor are listed in (21)-(22).

(21) Mother: *hine, ma ani osa?*
'there, what am I doing?'

Smadar: ***megida*** *et ha-shafan*.[10]
megida-S:FEM:PRES ACC the bunny

Mother: *ve ma ha-shafan ose?*
'And what does the bunny do?'

Smadar: ***mangid*** *et acmo.*
mangid-S:MASC:PRES ACC himself

Mother: *ve ma ani osa im ha-barvaz?*
'And what am I doing with the duck?'

Smadar: ***mangida*** *oto.*
mangida-S:FEM:PRES him [2;0]

(22) Leor: *ma savta* ***mebabashet****?*[11]
what grandma ***mebabashet***-S:FEM:PRES

Aunt: *ma savta ma?*
what grandma what?
'Grandma does what?'

Leor: ***mibabeshet****.*
mibabeshet-S:FEM:PRES

Aunt: *savta mitlabeshet? savta loh mitlabeshet.*
Grandma (is) getting dressed-FEM?
'Grandma (is) not getting-dressed-FEM'

Leor: *savta* ***mibaybaesh.***
grandma ***mibaybaesh***-S:MASC:PRES [2;3]

The verb-forms in (21) are derived from the common child language forms *tagidi* 'say-2S:FEM:FUTIMP' and *lehagid* 'say-INF', and overextend existing verb-forms to fill a morphologically defective paradigm. Lexically, except in the future, imperative, and infinitive, a suppletive form is used for *say* (*amr1* 'say', or *spr3* 'tell'). Phonologically, the root initial *n* (which occurs in Smadar's *ma**n**gida*) does not, in fact, show up in any of the adult forms (cf. adult *nafal – yipol* vs. children's *nafal – yi**n**pol* 'fall down', *natati – nata**n**ti* 'give-1S:PAST', *esa – e**n**sa* 'go (by car)-1S:FUT'). In this example, Smadar uses the innovated verb *mangid* with a noun complement in post-verbal position, analogously with other transitive verbs in her lexicon, e.g., *ha-shafan mesaper sipur* 'The bunny tells a story'. The verb-form in (22), *mibaybaeshet,* is a blend of *mitlabeshet* 'gets dressed-FEM' and *mitbayeshet* 'is ashamed-FEM'. Leor uses this innovated verb with a noun

complement in subject position, analogously with other intransitive verbs from his already established repertoire, e.g., *aba mitgaleax* 'Daddy is shaving' [1;10].

These innovations demonstrate that children use novel or self-created verbs in familiar patterns, rather than invent new argument structures for them. This, in turn, suggests that children form these verb-argument combinations by applying a rule rather than by repeating rote-learned utterances.

3.3.2. Overextension Errors

In Hebrew, verbs are based on the integrated constructs of *consonantal root* and *affixal pattern* called *binyan* conjugations. The five major morphological patterns are shown in Table 9 for the root *k-t-b* 'write' (The capital C's mark the positions of the root consonants in each pattern).[12]

Table 9. Hebrew Verb-Patterns

Verb Pattern		*k-t-b*	*Gloss*
P1	*CaCaC*	*katav*	'write'
P2	*niCCaC*	*nixtav*	'be/get written'
P3	*CiCeC*	*kitev*	'captionize'
P4	*hitCaCeC*	*hitkatev*	'correspond'
P5	*hiCCiC*	*hixtiv*	'dictate'

Binyan patterns form the basis for morphological marking of predicate-argument relations like transitivity, causativity, passive vs. middle vs. active voice, reciprocality, reflexivity, and incohativity, so that acquisition of verb syntax and semantics involves command of a fixed set of morphological patterns (Berman 1985, 1993a,b). Thus, VAS alternations at the level of the sentence almost always entail morphological alternation at the level of the verb, marked by a shift in *binyan* assignment. For example, the root *r-k-d* 'dance' can be inserted into the P1 pattern to yield the intransitive verb *rakad* 'dance' or into the P5 pattern to yield the causative transitive verb *hirkid* 'make dance'.

Against this background, another type of evidence for children's rule-bound behavior occurs in the form of overextension errors, i.e., the use of intransitive verbs in transitive argument structures. This is illustrated below with the Hebrew particular phenomenon of morphological verb-pattern alternation for marking verb-transitivity (Berman 1980, 1982, 1986, 1993a,b). Examples (23)-(24) list overextension errors made by Leor, and examples (25)-(27) overextension errors made by Smadar.

(23) *ani epol otax*
I fall-1S:FUT you-2S:FEM
'I will fall you = I'll drop you' [2;8]

nopel otax
fall-S:MASC:PRES you-2S:FEM
'(I) fall you = I drop you' [2;8]

(24) *oxelet et Leori*
eat-S:FEM:PRES Leori
'eating Leori = feeding Leori' [2;8]

(25) *ani rokedet oto*
I dance-S:FEM:PRES him
'I'm dancing him = I'm making him dance' [2;2]

(26) *Aba herim oti ve ala oti*
Daddy pick-3S:MASC:PAST me and get-up-3S:MASC:PAST me
'Daddy picked me up and got me up = lifted me up' [2;2]

(27) *Miryam overet te kol ha-dapim*
Miryam cross-S:FEM:PRES all the pages
Miryam crosses all the pages = turns all the pages over [2;3]

Example (23) shows that Leor overextends the use of the root *n-p-l* 'fall' in the intransitive P1 pattern to denote the causative action 'make-fall = drop'. This is highly ungrammatical, since, as noted, causativity is marked in Hebrew by inserting a root into the P5 pattern (cf. P5 *hipil*). Example (24) shows that Leor overextends the use of *a-k-l* 'eat' in the P1 pattern to denote the causative action 'feed' (cf. P5 *ma'axila*), again yielding an ungrammatical sentence. Similarly, examples (25)-(27) show that Smadar uses the roots *r-q-d* 'dance', *i-l-y* 'go up' and *i-b-r* 'cross' in the intransitive P1 pattern to denote the causative actions 'make-dance' (cf. P5 *markida*), 'lift up' (cf. *he'ela*), and 'turn over' (cf. P5 *ma'avira*), respectively, producing overextension errors. Similar findings were attested for Hagar, as well as for other Hebrew-speaking children studied by Berman (1980, 1982, 1985, 1993a,b). These findings suggest that children already know they need to use a transitive argument structure to denote causativity, and they try to apply this knowledge across-the-board. What they still do not know is how to encode causativity through morphology, as required in Hebrew (i.e., by verb-pattern alternations).

Hebrew-speaking children come to realize that the grammar of their language requires morphological marking of argument structure alternations typically from around age 3, after simple clause structure is established (Berman 1985, 1993a). This is illustrated in example (28) with data from Leor.

(28) *hipalti otax*
make-fall-1S:PAST you-2S:FEM
'(I) dropped you' [2;10]

ha-katar hipil ota
the locomotive make-fall-3S:MASC:PAST her
'The locomotive dropped her' [2;11]

ani apil lax me-ha-rosh
I make-fall-1S:FUT to-you-2S:FEM from-the-head
'I will make-fall to-you from-the-head = I'll drop (something) off your head' [3;0]

Example (28) shows that only at around age 3 Leor starts to alternate the familiar P1 pattern (which he initially used with the root *n-p-l*) with the P5 pattern correctly to yield the causative *hipil* 'drop' (cf. example [23] above). In sum, I have argued that the final phase of VAS acquisition is characterized by application of rules that children have formulated on the basis of previous experience with linguistic input. Two types of evidence supported this claim – the use of innovated verbs in familiar argument structure patterns, and the occurrence of overextension errors. The latter type of evidence further suggests, that it does not suffice to acquire general argument structure patterns to master VAS in a particular language. Language specific mechanisms, like Hebrew verb-pattern alternations for marking predicate-argument relations must also be acquired to complete the process.

4. CONCLUSION

The present study has laid out a developmental account of VAS acquisition based on analysis of longitudinal data from child Hebrew. Three developmental phases emerge from the data, as follows. Children start out with a limited range of bare verbs, each in a particular morphological form, initially determined by frequency in input to the child, and by the communicative function of each verb, e.g., transfer verbs like 'give' and 'bring' are used in the imperative form to make requests (see sections 3.1.1, 3.1.2, 3.1.3). Along with these verbs or soon afterwards, children start

producing verb-specific "holophrases", or frozen verb-argument chunks (*Input-based* phase).

Having heard and (presumably) stored a range of verbs from the input, children seem to engage in distributional analyses to help them come up with approximations of argument structures for these verbs. They first show uncertainty about argument position, as suggested by the early use of "groping patterns" (Braine 1976). Later, these are replaced by formulaic use of certain [V + N] schemas in repeated contexts, demonstrating positional consistency of arguments. Children, then, gradually increase the variety of lexical items (e.g., nouns) in a designated argument slot, as well as the range of different argument types used with each verb (*Bottom-up Construction of Generalizations*). Initially, these early generalizations emerge for a limited set of frequently used verbs, e.g., *rcy1* 'want', *ntn1* 'give', *isy1* 'make/do', *gmr1* 'finish' (Berman and Armon-Lotem 1996; Uziel-Karl 1997, 2001).

As more and more verbs interact with more and more argument slots to achieve a "critical mass" (Marchman and Bates 1994; Plunkett and Marchman 1993), knowledge of VAS becomes increasingly top-down and constructionist, so that children associate new verbs that enter their lexicon with argument structures from their already established repertoire (*from Generalizations to Rules*). Thus, the general progression is bottom-up to top-down, from specific items to linear stringing of constructions in which these items occur to hierarchical structures, from most concrete to most abstract, from item-specific to construction-based.

Several issues remain to be explored by future research. It was noted above, that for each new level of knowledge to be achieved, it must first attain a "critical mass" as input. This may take several forms – a large enough number of tokens of a particular verb, of verb-types that enter into a given "construction", or of verbs with different valency values. An important issue is whether all of these are sufficient and/or necessary requirements for achieving mastery of VAS in a particular language. Another major issue relates to the transitions between developmental phases. In particular, what factors trigger these transitions, whether these factors vary across development, and if so, what is the nature of that variation?

Finally, underlying the present account is the assumption that since children need to acquire a complex array of communicative knowledge on various levels, it is reasonable that they will use bits of whatever they know about linguistic form and language use to learn more. From this perspective, the language learner is viewed as an active participant in the acquisition process. However, it is still unclear to what cues in the input children are more attentive at each developmental phase, and to what extent language comprehension affects production of verb argument structure.

Well-motivated answers to these questions lie beyond the scope of this study, and would require large-scale longitudinal sampling, supplemented by structured-elicitations and experimental designs, as well as examination of developmental data from typologically different languages.

SIGAL UZIEL-KARL
Kibbutzim College of Education, Tel Aviv

ACKNOWLEDGEMENTS

This paper is dedicated to my late grandfather Mois Amarigllio.

5. NOTES

[1] The following abbreviations are used throughout the paper: VAS – Verb Argument Structure, V – Verb, VI – Verb + Indirect (dative) object, VO – verb + direct object, SVO – Subject, Verb, Object; S – Singular, P – Plural; MASC – Masculine, FEM – Feminine; FUT – Future, FUTIMP – Future-Imperative, IMP – Imperative, INF – Infinitive, PRES – Present, PAST – Past; US – Unspecified, ACC – Accusative; MLU – Mean Length of Utterance.

[2] Armon-Lotem (this volume, section 4.1) claims that Hebrew-speaking children start using gender marking equally for feminine and masculine already at age 1;6. Her claim appears to contradict the claim made in the present study regarding the early use of *rcy1* 'want'. This apparent contradiction can be accounted for as follows. First, Armon-Lotem based her generalizations on verbs in the imperative, rather than in the present tense. In the early phases of acquisition, this makes a difference, since as it was argued here (section 3.1), each verb is initially acquired individually. Secondly, the verb *rcy1* 'want' has been noted to form an exception to other early verbs in its patterns of use (Ninio 1999; Uziel-Karl 2001). Finally, the differences between the present paper and Armon-Lotem's paper may be due to methodological differences in the definitions and criteria for "productivity" and "acquisition".

[3] The text in brackets lists the standard adult form.

[4] In the Table, the numbers in brackets indicate the number of tokens of each verb-form in the database.

[5] In the Table, *A* stands for *Adult* and *C* for *Child*.

[6] The large gap between the age at which Smadar starts producing the attested "groping patterns" and the age at which Lior starts producing them can be attributed to individual differences in the girls' overall linguistic development as measured by MLU. Thus, while Smadar's MLU between ages 1;6 - 1;7 ranges between 1.5 - 2, Lior matches this MLU range only between ages 1;8 - 2;1. It is hard to determine the exact age at which "groping patterns" first occurred in the speech of Leor and Hagar, since data collection for these two children started only at age 1;9 and 1;7, respectively. It should be noted, however, that both children had a prolonged period at which their MLU = 2. A more detailed comparison of the four children in the sample is provided in Uziel-Karl (2001).

[7] *Sorya* is a nonsense word.

[8] I wish to thank an anonymous reviewer for bringing this work to my attention.

[9] Early morphological development is illustrated with examples from Lior and Smadar, since data collection for these two girls started much earlier than for the other two children in my corpuse (Leor and Hagar), and as a result, it is much easier to detect developmental trends in their data.

[10] The meanings of *mangid* and *mangida* are not entirely clear from the conversational context. The way Smadar uses these innovated verbs, they seem to denote some kind of action performed by an Agent on a Patient, or by a Patient on itself.

[11] The innovated verb *mibabeshet* seems like a mispronunciation of the already existing verb *mitlabeshet* 'get-dressed-S:FEM:PRES'.

[12] The stops /k/ and /b/ alternate with the spirants /x/ and /v/ in various morphological contexts.

6. REFERENCES

Akhtar, Nameera. "Acquiring Basic Word Order: Evidence of Data-driven Learning of Syntactic Structure." *Journal of Child Language* 26 (1999): 339-356.

Akhtar, Nameera and Michael Tomasello. "Young Children's Productivity with Word Order and Verb Morphology." *Developmental Psychology* 33 (1997): 952-965.

Allen, Shanley E. M. *Aspects of Argument Structure Acquisition in Inuktitut*. Amsterdam: John Benjamins, 1996.

Ariel, Mira. *Accessing Noun Pharse Anteceedents*. London: Routledge, 1991.

Armon-Lotem, Sharon. *Parametric Approach to Functional Heads and the Acquisition of Hebrew*. Unpublished doctoral dissertation, Tel-Aviv University, 1997.

Armon-Lotem, Sharon. "Subjects Use and the Acquisition of Verbal Agreement in Herbrew". In Natalia Gagarina and Insa Gulzow (eds.), *The Acquisition of Verbs and their Grammar: The Effect of Particular Languages*. Springer, (to appear).

Barrett, Martin. "The Holophrastic Hypothesis: Conceptual and Empirical Issues." *Cognition* 11 (1982): 47-76.

Bates, Elisabeth and Brian MacWhinney. "Competition, Variation and Language Learning." In Brian MacWhinney (ed.), *Mechanisms of language acquisition*. Hillsdale, NJ: Erlbaum, 1987: 157-193.

Bates, Elisabeth and Brian MacWhinney. "Functionalism and the Competition Model." In Brian MacWhinney and Elisabeth Bates (eds.), *The Crosslinguistic Study of Sentence Processing.* New York: Cambridge University Press, 1989.

Berman, Ruth A. "The Case of an (S)VO Language: Subjectless Constructions in Modern Hebrew." *Language* 56 (1980): 759-776.

Berman, Ruth A. "Verb-pattern Alternation: The Interface of Morphology, Syntax and Semantics in Hebrew Child Language." *Journal of Child Language* 9 (1982): 169-191.

Berman, Ruth A. "Acquisition of Hebrew." In Dan I. Slobin (ed.), *The Cross-linguistic Study of Language Acquisition,* Vol. 1. Hillsdale, NJ: Erlbaum, 1985: 255-371.

Berman, Ruth A. "A Step-by-Step Model of Language Learning." In Iris Levin (ed.), *Stage and Structure: Reporting the Debate.* Norwood, NJ: Ablex, 1986: 191-219.

Berman, Ruth A. "Developmental Perspectives on Transitivity: A Confluence of Cues." In Yonata Levy (ed.), *Other Children, Other Languages: Issues in the Theory of Language Acquisition.* Hillsdale, NJ: Erlbaum, 1993a: 18

Berman, Ruth A. "Marking of Verb Transitivity by Hebrew-speaking Children." *Journal of Child Language* 20 (1993b): 641-669.

Berman, Ruth A. and Sharon Armon-Lotem. "How Grammatical are Early Verbs?" In Claire Martinot (ed.), *Actes du Colloque International sur L'acquisition de la Syntaxe en Langue Maternelle et en Langue Etrangère.* Paris: Les Belles Lettres, 1996: 17-59.

Berman, Ruth A. and Jurgen Weissenborn. "Acquisition of Word Order: A Crosslinguistic Study." Final Report submitted to the German-Israeli Foundation for Scientific Research and Development (G.I.F.), Bonn and Jerusalem, 1991.

Bittner, Dagmar, Wolfgang U. Dressler and Marianne Kilani-Schoch (eds.), *First Verbs: On the Way to Mini-Paradigms*. *Zas Papers in Linguistics* 18, 2000.

Bowerman, Melissa. *Early Syntactic Development: A Cross-linguistic Study with Special Reference to* Finnish. Cambridge: Cambridge University Press, 1973.

Bowerman, Melissa. "Reorganizational Processes in Lexical and Syntactic Development." In Lila Gleitman and Eric Wanner (eds.), *Language Acquisition: The State of the Art.* New York: Cambridge University Press, 1982: 319-346.

Braine, Martin D. S. "Children's First Word Combinations." *Monographs of the Society for Research in Child Development* 41.1. Chicago: University of Chicago Press, 1976.

Bruner, Jerome S. *Child's Talk.* New York: Norton, 1983.

Bybee, Joan L. and Dan I. Slobin. "Rules and Schemes in Development and Use of the English Tense." *Language* 58 (1982): 265-289.

Clark, Eve V. *The Lexicon in Acquisition*. Cambridge: Cambridge University Press, 1993.

Clark, Eve V. "Language Acquisition: The Lexicon and Syntax." In Jon Miller and Peter Eimas (eds.), *Speech, Language and Communication.* New York: Academic Press, 1995: 303-337.

Ferdinand, Astrid. *The Development of Functional Categories: The Acquisition of the Subject in French.* Doctoral dissertation. HIL, Leiden University, 1997.

Gathercole, Virginia, Eugenia Sebastian, and Pilar Soto. "The Early Acquisition of Spanish Verbal Morphology: Across-the-board or Piecemeal Knowledge?" *International Journal of Bilingualism* 3 (1999): 133-182.

Gleitman, Lila R. "The Structural Sources of Verb Meaning." *Language Acquisition* 1 (1990): 3-55.

Hirsh-Pasek, Kathy and Roberta M. Golinkoff. *The Origins of Grammar: Evidence from Early Child Comprehension.* Cambridge, MA: MIT Press, 1996.

Ingram, David, Anne Welti and Christine Priem. "The Early Stages of Verb Acquisition in Spanish, German and English". In Natalia Gagarina and Insa Gulzow (eds.), *The Acquisition of Verbs and their Grammar: The Effect of Particular Languages*. Springer, (to appear).

Kilani-Schoch, Marianne and Wolfgang U. Dressler. "The Emergence of Verb Paradigms in Two French Corpora as an Illustration of General Problems of Pre- and Proto-Morphology." Poster Presented at the *9th International Morphology Meeting* (2000).

Lieven, Elena, Julian Pine, and Gillian Baldwin. "Lexically-based Learning and Early Grammatical Development." *Journal of Child Language* 24 (1997): 187-220.

MacWhinney, Brian. "Processing a First Language: The Acquisition of Morphophonology." *Monographs of the Society for Research in Child Development* 43. Chicago: University of Chicago Press, 1978.

MacWhinney, Brian. "Basic Processes in Syntactic Acquisition." In Stan Kuczaj (ed.), *Language Development: Syntax and Semantics.* Hillsdale, NJ: Erlbaum, 1982: 73-136.

MacWhinney, Brian. *The CHILDES Project: Tools for Analyzing Talk.* Hillsdale, NJ: Erlbaum, 1995.

Marchman, Virginia and Elisabeth Bates. "Continuity in Lexical and Morphological Development: A Test of the Critical Mass Hypothesis." *Journal of Child Language* 21 (1994): 339-366.

Maratsos, Michael P. and Mary Anne Chalkley. "The Internal Language of Children's Syntax: The Ontogenesis and Representation of Syntactic Categories." *Children's Language* 2 (1981): 127-214.

Ninio, Anat. "On formal Grammatical Categories in Early Child Language." In Yonata Levy, Izchack M. Schlesinger, and Martin D. S. Braine (eds.), *Categories and Processes in Language Acquisition.* Hillsdale, NJ: Erlbaum, 1988: 11-30.

Ninio, Anat. "Pathbreaking Verbs in Syntactic Development and the Question of Prototypical Transitivity." *Journal of Child Language* 26 (1999): 619-653.

Ninio, Anat and Catherine E. Snow. "Language Acquisition through Language Use: The Functional Sources of Children's Early Utterances." In Yonata Levy, Izchack M. Schlesinger, and Martin D. S. Braine (eds.), *Categories and Processes in Language Acquisition.* Hillsdale, NJ: Erlbaum, 1988: 11-30.

Pinker, Steven. *Language Learnability and Language Development.* Cambridge, MA: Harvard University Press, 1984.

Pinker, Steven. *Learnability and Cognition: The Acquisition of Argument Structure.* Cambridge, MA: MIT Press/Bradford Books, 1989.

Pizzuto, Elena and Christina Caselli. "The Acquisition of Italian Morphology." *Journal of Child Language* 19 (1992): 491-557.

Pizzuto, Elena, and Christina Caselli. "The Acquisition of Italian Verb Morphology in a Cross-linguistic Perspective." In Yonata Levy (ed.), *Other Children, Other Languages: Issues in the Theory of Language Acquisition.* Hillsdale, NJ: Erlbaum, 1993: 137-187.

Plunkett, Kim and Virginia Marchman. "From Rote Learning to System Building." *Cognition* 48 (1993): 21-69.

Schlesinger, Izchak M. "The Origin of Relational Categories." In Yonata Levy, Izchack M. Schlesinger, and Martin D. S. Braine (eds.), *Categories and Processes in Language Acquisition.* Hillsdale, NJ: Erlbaum, 1988: 121-178.

Stern, Naftali. "The 'et' Verbs in Israeli Hebrew." *Hebrew Computational Linguistics* 15 (1979): 28-57. [in Hebrew]

Stern, Naftali. "Three-place Verbs Containing the 'et' Complement and Others." *Hebrew Computational Linguistics* 17 (1981): 46-61. [in Hebrew]

Tomasello, Michael. *First verbs: A Case Study of Early Lexical Development*. Cambridge University Press, 1992

Tomasello, Michael. "First Steps toward a Usage-based Theory of Language Acquisition." *Cognitive Linguistics* 11.1/2 (2000a): 61-82.

Tomasello, Michael. "Do Young Children Have Adult Syntactic Competence." *Cognition* 74 (2000b): 209-253.

Tomasello, Michael and Patricia Brooks. "Young Children's Earliest Transitive and Intransitive Constructions." *Cognitive Linguistics* 9 (1998): 379-395.

Uziel-Karl, Sigal. "How General is Verb Argument Structure? The Developmental History of Some Early Verbs in Hebrew." *Proceedings of GALA* (1997): 172-177.

Uziel-Karl, Sigal. "Children's Verb Lexicon." Proceedings of the *30th Child Language Research Forum Conference* (2000): 41-49.

Uziel-Karl, Sigal. *A Multidimensional Perspective on the Acquisition of Verb Argument Structure.* Unpublished doctoral dissertation, Tel-Aviv University, 2001.

Uziel-Karl, Sigal. "Acquisition of Verb Paradigms across Verb Classes: Evidence from Child Hebrew." Paper presented at the *Pre- and Proto-Morphology Workshop*, Vienna, 2004.

Uziel-Karl, Signal and Ruth A. Berman. "Where's Ellipsis? Whether and Why There are Missing Arguments in Hebrew Child Language". *Linguistics* 38 (2000): 457-482.

SHARON ARMON-LOTEM

SUBJECT USE AND THE ACQUISITION OF VERBAL AGREEMENT IN HEBREW*

Abstract. This paper presents an analysis of the early use of verbal morphology and overt subjects in Hebrew before the age of two, and the way in which the two processes interact, leading to some agreement mismatches. In Hebrew (Armon-Lotem 1996a), the acquisition of aspect and tense morphology is interrelated with the acquisition of agreement morphology, in a way that supports proposals for a bottom-up acquisition of phrase markers (Vainikka and Young-Scholten 1996). Specifically, children's use of verbs in an aspectually limited manner (e.g., telic verbs only with past tense morphology and atelic verbs only with present tense morphology) is followed by the use of agreement markers for gender and number, then by the appropriate use of tense morphology and finally by the use of agreement markers for person. In subject position, children use bare nouns and proper names as overt subjects only after they use aspect, but before they have subject verb agreement in gender and number. Similarly, they use pronouns after they use tense and before they use person agreement. This order results in agreement mismatches. In a framework where heads are obligatory, specifiers are optional, and agreement is not a head but rather the mere syntactic relation between the two (Chomsky 1995), this head > spec > agreement > head > spec > agreement order is predicted. Specifically, assuming that there are no Agr nodes, Asp is specified with gender and number spec-features and T is specified with person spec-features. The paper addresses the question which bootstrapping strategies children use in acquiring these features. Children first use the Aspect head, next the appropriate specifier, and then apply the syntactic operation of spec-head agreement resulting in agreement in gender and number. Later, when the semantically motivated functional head T is acquired, pronouns emerge in its specifier, and spec-head agreement in person applies. Thus, the interaction between the early use of verbal morphology and the early use of subjects can be accounted for by the "no Agr nodes" approach.

1. INTRODUCTION

Within the generative framework, the finite verb is the grammatical center of the sentence, since IP and CP are the functional projections of V (Grimshaw 1991). Thus, we expect to find interrelationship between the acquisition of finite verbs and that of the morphosyntactic features of sentential constituents which depend on the verb. This paper studies the interrelationship between the acquisition of verbal morphology and the use of overt subjects in Hebrew. The nature of Hebrew, with its rich verbal morphology which clearly distinguishes person from gender and number, and agreement from tense, is reflected in a fine grained picture of the developmental process, which is less apparent in other languages. The paper presents an analysis of the early acquisition of verbal morphology before the age of two, when children move

N. Gagarina and I. Gülzow (eds.), The Acquisition of Verbs and their Grammar: The Effect of Particular Languages, 45–67.

from early verbal combinations to the last stage of verbal acquisition: early paradigm acquisition (Ingram, Welti and Preim, this volume). The analysis proposed in this paper can serve as a possible outline of the substages of this last stage of verbal acquisition. This analysis is followed by a discussion of the way in which the use of verbal morphology and the use of overt subjects are interrelated, yielding some agreement mismatches.

It is argued that the agreement mismatches discussed below are not arbitrary, but rather mirror a gradual deployment of the full range of functional heads, in a bottom-up order (Armon-Lotem 1996a). This kind of analysis has already been proposed for the nominal phrase (NP). Armon-Lotem (1998) shows that in the acquisition of the various component of NP, deviations from the adult pattern, e.g., in word order, are due to a gradual deployment of the nominal system. The question addressed in this paper is which bootstrapping strategies are used in order to specify the features of the functional heads which expands the VP, enabling their gradual deployment.

Moreover, it is suggested in this paper that the "economy of derivation" (Speas 1994) which yields a "no-Agr nodes" approach (Chomsky 1995) actually predicts the possibility of such agreement mismatches, thus being more explanatory adequate. The "economy of derivation" is an attempt to limit the possible range of functional heads by defining which morpho-syntactic features are overt representations of functional nodes. It is proposed that the inventory of functional heads should ideally be limited to semantically motivated heads. Thus, tense or definiteness, being encored in the semantics, must also be manifested syntactically by a functional head in order to enable one-to-one mapping of the two systems. On the other hand, agreement or case, which uniquely mark syntactic relations, need not correlate with one specific functional head, but rather be defined as rules on the relations between different semantically motivated nodes. These are the assumptions behind Chomsky's (1994) "no-Agr nodes" approach, that is, the proposal that there should be no agreement nodes, since spec-head agreement is a purely syntactic notion. By this approach only Tense and Aspect are valid functional categories at the IP level.

Finally, it is shown in this paper how children's use of verbal morphology and overt subjects complies with locality principles (Chomsky 1995). The locality principle intends to determine which components are minimally obligatory in any phrase marker. The assumption is that heads are obligatory being the core of the phrase and complements are obligatory by projection from the lexicon. A head and its complement form the most local domain. Specifiers, on the other hand, are optional, and are therefore in the local, but not most local, domain. This assumption has already been employed in studies of language acquisition in order to explain the delayed use of specifiers (see also Clahsen, Eisenbeiss and Vainikka 1994).

After a short overview of Hebrew verbal morphology, the paper will discuss the order of acquisition of the verbal system in Hebrew focusing on verbal inflections as well as children's use of subjects. These two processes and the way they interrelate will raise a few questions, which will finally be discussed under the assumption that children are "minimalists".

2. HEBREW VERBAL MORPHOLOGY

Hebrew verbs are constructed in one of five morphological patterns (conjugations). Each conjugation associates prefixes as well as interdigited vowels with a set of root consonants. Some of these patterns are associated with particular functions, such as reflexivity or causativity. However, there are numerous lexical gaps and semantic inconsistencies between form and function in this system, setting it in the domain of word-formation (derivational morphology) rather than grammatical inflection.

Hebrew verbs are also inflected for five tense/mood categories: infinitives, imperatives, present, past and future, as shown in Table 1 for the root *g.m.r* 'finish':

Table 1. The Hebrew tense/mood paradigm

Root	*Infinitive*	*Past*	*Present*	*Future*	*Imperative*
g.m.r.	*ligmor*	*gamar*	*gomer*	*yigmor*	*gmor*
	To finish	Finished	Finishing	Will finish	Finish!

As can be seen in Table 1, infinitives are marked by a prefixal *l-*. This prefix takes a different vowel depending on the nature of the following syllable (*li-*, *la-*, or *le-*). The past, the present, and the imperative are distinguished by the interdigited vowels, and the future is distinguished from the imperative by an agreement prefix. The imperative form in Table 1 is of the historically, normative type, based on future-tense stems. Children generally use the forms common in everyday colloquial usage, the future stem either with or without the 2nd-person *t-* prefix, depending on the vowel-pattern of the conjugation (Berman 1985; Bolozky 1979).

All finite forms and the imperative which is derived from the future tense are marked for agreement. Verbs in the imperative and the present tense, which is also the participial form, agree with their subjects in number and gender. This yields four forms per verb. For example, in the present tense, masculine singular has a zero morpheme, feminine singular is marked with unstressed *-et*, masculine plural with stressed *-im*, and feminine plural with stressed *-ot.* Verbs in past tense and future tense agree with their subjects in number, gender and person. This yields up to nine forms per verb. For

example, past tense affixes are added to the masculine singular stem: *-ti* 1st singular, *-ta* 2nd masculine singular, *-t* 2nd feminine singular, *-a* 3rd feminine singular, *-nu* 1st plural, *-tem* 2nd masculine plural, *-ten* 2nd feminine plural and *-u* 3rd person. Thus, the child has to deal with over thirty different forms for each verb in the active mode and a similar number in the passive.

This variety of inflections led to Shlonsky's (1995) proposal of an elaborated structure for the functional system in Hebrew phrase markers. As in other languages (Kayne 1994), it is assumed that Infl, the head where all inflectional features are marked and checked, is separated into agreement, tense and aspectual categories, each with its unique functional heads. Agreement itself, however, is divided into two: participial agreement (AgrPrt) and subject agreement (AgrS)[1]. It is AgrPrt, where gender and number participial features are checked and AgrS, where person features are checked (Shlonsky 1995; Siloni 1994). The functional heads are ordered as illustrated in Figure 1:

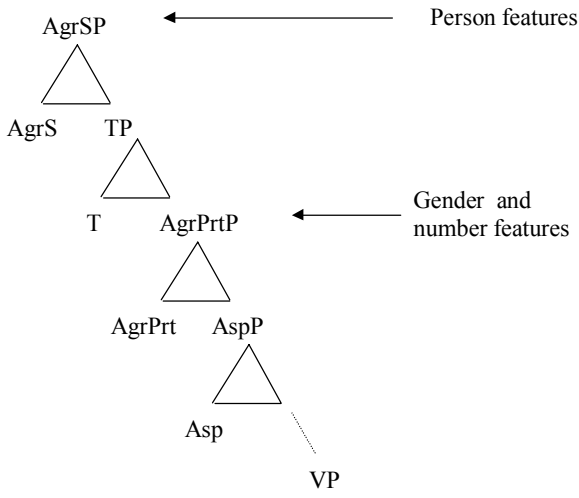


Figure 1. The Hebrew phrase marker

This phrase-marker is used by Shlonsky to account for various syntactic phenomena in Hebrew, including the distribution of null-subjects. In this paper, the focus is on the way in which children acquire this phrase-marker.

3. THE CORPUS

The data analyzed in this paper come from 3 Hebrew-speaking children, two girls, Lior (1;5 - 3 :1) and Smadar (1;4 - 2;3), and one boy, Leor (1;9 - 3;0). All three children come from middle class families who live in Tel-Aviv or its suburbs, in which at least one parent has obtained higher education. Leor and Lior are first children, with no younger siblings (Lior's baby brother is born during the period of recording), and Smadar is the youngest in her family, having two older (preschool and school age) siblings. The data were collected by the Tel-Aviv University Language Acquisition Project, as part of a crosslinguistic study of early word order directed by Ruth Berman (Tel-Aviv University) and Juergen Weissenborn (University of Potsdam). The data are supplemented by data from Ravid Corpus in CHILDES.

Data for the three children were taken from weekly longitudinal recordings of the children's spontaneous speech output in interaction with an adult (their parents, and in Leor's case, his aunt), starting with their first word combinations. Leor's aunt and Smadar's mother were linguistics students at the time of recording. Table 2 below gives the gender and age of the children in the study and the number of recording and number of child utterances per recording as well as MLU range for the individual children:

Table 2. Breakdown of 3 longitudinal samples analyzed for the emergence of grammar

Child	*Age range*	*Number of transcripts*	*Range of child utterances across transcripts*	*Range of MLU across transcripts*
Leor, boy	1;9.0 - 2;0.30	22	56 - 316	2.11 - 3.38
Lior, girl	1;5.19 - 1;11.30	82	68 - 256	1.13 - 2.22
Smadar, girl	1;4.14 - 1;11.18	21	76 - 280	1.34 - 3.57

The MLU and age differences between the three children could have called for individual analyses, which were done initially, indicating that all three children can in fact be treated together. Crucially, our initial analyses implied that MLU is not indicative of early verb use, as far as lexical, semantic and verb pattern distributions are concerned. Moreover, there were no differences in these aspects between the first 20 verbs used prior to word combinations as documented for three diary studies and the verbs used in early verb combinations and simple sentences by these three children (Armon-Lotem and Berman, 2003). Inflectional markings, on the other hand, seem to be MLU related, showing changes over time. No correlation was

found, however, between MLU and the onset of use of verbal inflection and all three children follow the same course of development, no matter when they started using it, making it possible to discuss the findings for all three together.

The data is analysed separately for each inflectional category – gender and number, tense, and person. The distribution of tenses is divided into six categories: the five inflectional categories of mood/tense and a sixth category for unclear forms. Unclear forms are typically 'stemlike' stripped forms: they contain some or all of the root consonants of the verb and also some or all of the interdigited vowels associated with the particular verb conjugation, but no inflectional affixes which can determine their tense/mood category (either as irrealis or as finite), e.g., *pes* for either infinitive *letapes* 'to climb', present tense *metapes* 'climb, ms', *metapeset* 'climb, fm', future tense *yetapes* '(will) climb', or mispronounced forms in which the tense is unclear. The other five tense/mood categories are identified either by prefixes and suffixes, or by the internal vowel pattern associated with that category.

The findings for each category are presented by the percentage of verb forms inflected for that category out of the total number of verb forms. The term verb form, rather than type or lexeme, is used here for the combination of verb and inflections. The same verb type can have a few verb forms, i.e., could show up with different inflections. Tokens can be counted for each verb form separately, and if necessary can then be calculated for each type. A high verb type – verb form ratio can be used as a measure of productivity. This decision was made due to the nature of the data at the earlier phases of verb acquisition. When verbs are initially used, each verb form is assumed to be unanalyzed, making it impossible to treat two different inflections of the same verb as different tokens.

4. FINDINGS

4.1. The acquisition of verbal inflection

Studying the acquisition of the verbal system in Hebrew and focusing on the twenty earliest verbs of six Hebrew-speaking children prior to their second birthday, Berman and Armon-Lotem (1997) show that early verb usage is largely rote-learned and morphologically unanalyzed. Phonetically, most forms are reduced, non-adult forms, e.g., *ce* for *roce* 'want', *ito* for *lishtot* 'to drink'. Syntactically, over 80% of the verbal forms look like infinitives without the infinitival morpheme *le-* 'to' or like imperatives. Tense morphology is initially used to mark lexical aspect rather than tense: past tense forms are used only with perfective telic actions, e.g., *fal* for *nafal* 'fell' and present tense forms are used for durative atelic actions, e.g., *xel* for *oxel*

'eating'. Morphologically, all non-infinitival forms are singular, with some nonproductive variation between masculine and feminine forms, e.g., *kxi* 'you, feminine, take' is used both for female and male addressees. Finally, there is no overt person morphology on the verbs.

Armon-Lotem (1996a) analyzes the acquisition of verbal morphology by toddlers aged 1;6 to 2;0. She found that initially children do not inflect most of their verbs using many unclear forms, which cannot be distinguished for tense. Nonetheless, as reported in Berman and Armon-Lotem (1997), past and present morphology is used selectively according to the lexical aspect of some verbs, though aspect is not marked morphologically in the target language. These verb forms in which tense morphology is used in a non-target-like manner to mark aspect rather then tense are the first inflected verb forms, used even before children start combining words. These verb forms on which tense morphology is used for aspect will be referred as aspectually limited verb forms.

The first target-like morphological knowledge children manifest, as they start combining words, is subject-verb agreement in gender (and sometimes number). This knowledge is applied in a limited fashion to third person forms in the present tense and second person in the imperative. Initially, gender agreement is used on the verb in an unanalysed formulaic rather than a productive manner. This formulaic use yields utterances in which the verb shows no agreement with the subject, for example, when the child is using a verb with feminine inflections for a masculine subject and vice versa. Soon, however, the child converges on the grammatical use of gender and starts using it productively. Productive use is assumed only when the same lexeme is used grammatically with different inflections and the same inflection is used grammatically with different lexemes.

Figure 2 gives the breakdown, in percentage out of the total number of verb forms, of gender and number inflections in the verb usage of one girl, Smadar, from 1;6 to 1;11. The findings presented in figure 2 include both grammatical and ungrammatical use of the inflections as discussed below.

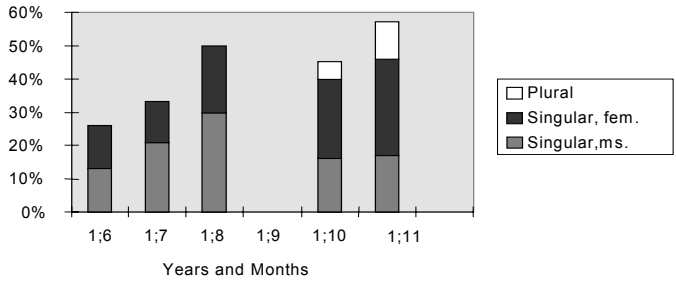


Figure 2. Breakdown of Gender and Number

As seen in the leftmost column of Figure 2, Smadar at age 1;6 starts using masculine and feminine gender marking in equal proportions. This is mostly in the imperative with one past form in the default 3rd person masculine. For the first few sessions, these inflections are used in a formulaic unanalysed rather than a productive manner, as evident from the luck of agreement with overt or covert subjects. Within the next month, Smadar's use of gender inflection is found in five imperative forms, where it is used productively, agreeing with the subject inferred from the context, as shown in (1):

(1) *sim / símí* 'put!' ms / fm,
kax / kxi 'take!' ms / fm,
vi / abí'i 'give (me)! bring!' ms / fm,
ten / tni 'give!' ms / fm and
shev / shvi 'sit!' ms / fm [Smadar 1;7]

The imperative, in which second person is the default, already seems adultlike once gender and number morphology are used.

Another girl, Lior, at age 1;6.13 shows similar alternations in the imperative, as well as a possible alternation in the present-tense form *pes / péset* 'climbs' ms / fm. A month later (Lior 1;7.16), in addition to using gender and number alternations in the imperative, e.g., *zuz/zuzi/zuzu* 'move!, sg,ms/sg,fm/pl.' [Lior 1;7], she also uses present-tense *boxé / boxá* 'cries' ms / fm. Since person is not marked in Hebrew present tense, children appear to use these forms properly at this point. However, the forms they use are still aspectually limited, i.e., the same verbs do not occur in the past form.

Figure 2 also points to a clear age-related difference in the acquisition of gender and number. While gender is used from the first sessions, number occurs only at age 1;10. These findings, however, were not the same for the three children; Leor, the boy, whose records start at 1;09, used plural forms from the very first sessions, though it has been used productively, i.e., in alternation with the singular, only two months later.

About a month after gender is used productively, a sharp increase in the use of past and present-tense forms is observed. For example, while Smadar at 1;7 uses only one form of the verb 'eat', at 1;10, the same verb shows the following distribution:

(2) *le'exol* 'to eat', *oxélet* 'eating' sg,fm, *axá(l)ti* 'I ate',
toxlí 'you (fm.) will eat', also colloquial imperative 'eat!'
[Smadar 1;10]

Figure 3 gives the breakdown, in percentage out of the total number of verb forms, for tense/mood categories in Smadar's verb usage at this point for the three major classes of forms she used: unclear forms, irrealis forms which combine infinitives and imperatives, and finite forms which combine present, past and future. Unclear forms are typically 'stemlike' stripped forms containing the root consonants and some or all of the interdigited vowels associated with the particular verb form, but no inflectional affixes which can determine their tense/mood category.

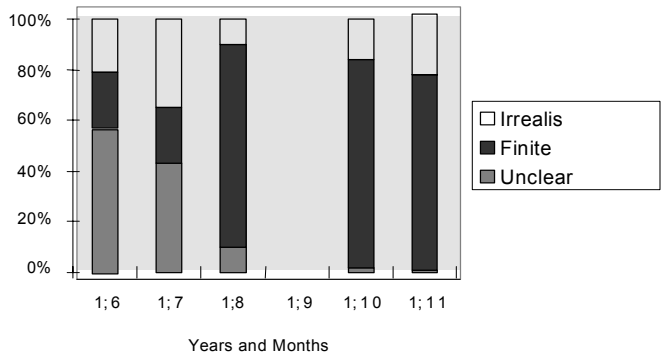


Figure 3. Breakdown of Tense/Mood

Figure 3 reveals an explosion of finite forms (mainly present and past) at age 1;8 to 1;10 months. A comparison of the use of unclear forms, irrealis forms (infinitives and imperatives), and finite forms (present, past and future) shows a remarkable change. This increase in tensed forms correlates with a decrease in the percentage of unclear forms. For Example, at age 1;7, tensed forms constitute only 22% of Smadar's verbal forms, and these are aspectually limited, whereas one month later, by age 1;8, tensed forms constitute 80% of her verbal forms (45 tensed forms out of 55 verb forms at 1;10), and unclear forms are very rare. That these verbs indicate tense rather than (lexical) aspect is evident by the use of the same lexeme both in past and present tense, as well as by the use of durative atelic verbs in the past, as shown in (3) and unaccusative telic (perfective) verbs in the present-tense or future (used as imperative), as shown in (4):

(3) *'hi boxa* 'she is crying', *hi baxta* 'she was crying' [Smadar 1;10]

(4) *sagarnu* 'we closed', *yisgor* 'he will close (used for imperative)' [Smadar 1;10]

Irrealis forms are used in an adultlike manner: imperative forms for requests and infinitival forms for questions and as replies to questions. Only a few infinitival forms (not more than 5% of all verb forms) are used ungrammatically in affirmative contexts, and can be labelled as root infinitives (Armon-Lotem 1996b).

The last inflections to be acquired during this period are first and second person markings. For the first few months, while agreement in gender and number, as well as tense, is acquired, children make only sporadic use of person agreement. Figure 4 gives the breakdown, in percentage out of the total number of verb forms for first, second and third person in Smadar's verb usage during the same period. Though verbs are inflected for number in Hebrew, for ease of presentation, we have combined the categories of number and person, neutralizing for number so that person applies to both singular and plural (the equivalent of both 'I' and 'we', and so on for 2nd and 3rd person, as well). This seemed legitimate since the children generally used only a few plural forms and the pattern is not substantially affected by separating out the category of number.

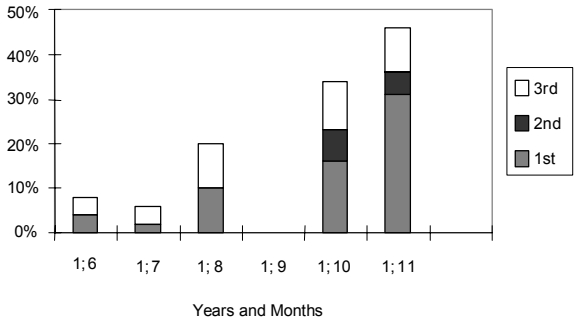


Figure 4. Breakdown of Person

Smadar used first person singular in the past, once per month over the first three months in what looks like rote-learned forms (Berman 1986). At 1;10, however, after she starts using different tenses for the same verbs, she moves into using both first and third person inflections, as in (5), with the same verb in a productive manner. Alternations between first and second person are found a month later.

(5) *asiti* 'I did', *asinu* 'we did', *asta* 'she did' [Smadar 1;10]

Lior, on the other hand, used no person morphology for the first two month, apart from one verb in second person in the past tense: *asit* 'made' meaning 'went potty'. Over the next two months, she added one more verb in second person *(na)falt* 'fell'. This order could look contradictory to the previous findings presented here, suggesting that second person is used before first person, except that Lior was not using these forms in a target-like manner. Rather, Lior uses *asit* and *(na)falt*, both of which have second person morphology to refer to herself in a first person sense. This changes by age 1;11, when person becomes a robust phenomenon which is used productively, and second person is used with the proper referent, see (6):

(6) *gamárti* 'I finished', *gamárta* 'you (ms.) finished',
gamart 'you (fm.) finished', *gamárnu* 'we finished'
[Lior 1;11]

By acquiring person inflection, children complete the acquisition of the inflectional verbal morphology of Hebrew. There are many irregularities in the verb forms that they still have to master, but none of them involves different inflections.

Figure 5 shows that the order of acquisition of the three inflectional categories – gender/number, tense/mood and person – is the same for all three children. For each of the children gender marking is the first to emerge; followed by an explosion of tensed forms marking tense/mood distinctions; while (past tense) suffixes and (future tense) prefixes marking first person in contrast with second and third person are the last to be acquired.

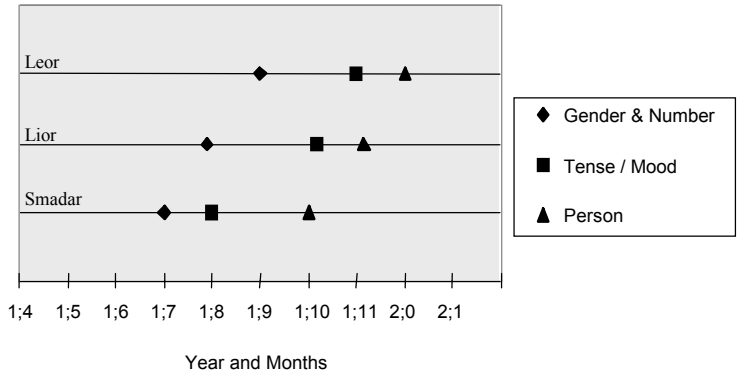


Figure 5. Age of Acquisition of Agreement and Tense/Mood Inflection

This order of acquisition raises a couple of questions. The first question is why children start by using tense inflections to mark lexical aspect rather than tense, that is, why they start with aspect even when it does not exist in the input (the adult language). The second question is why inflections are acquired in this order. Since the use of agreement morphology depends on the nature of the subject, the use of subjects in child Hebrew needs to be addressed before we turn to answer these questions.

4.2. The use of subjects in child Hebrew

Hebrew is a partial null-subject language, since *pro* is licensed only with two tenses: the past and the future (Berman 1978; Borer 1984). Referential *pro* is never licensed with the present participial form, though in plural masculine, this form licenses pro_{arb}, as shown in (7):

(7) a. *ba-kayic* *hem/*pro* *holxim* *la-yam*
in-the-summer they go to-the-beach
'In the summer, they go to the beach'

b. *ba-kayic* *pro* $_{arb}$ */*hem* *holxim* *la-yam*
in-the-summer they go to-the-beach
'In the summer, people go to the beach'

Moreover, even in the past and the future, in spoken Hebrew, *pro* is not licensed in all persons. It has been argued that in the past tense, *pro* is licensed only for first and second person, both singular and plural, but not for third person yielding the paradigm in Table 3 for the verb *halax* 'to go':

Table 3. The Paradigm for pro in Hebrew Past Tense

Person	*Number* / *Gender*	*Singular*	*Plural*
First		*Ani/pro halaxti* I went	*Anaxnu/pro halaxnu* We went
Second	Masculine	*Ata/pro halaxta* You went	*Atem/pro halaxtem* You went
	Feminine	*At/pro halaxt* You went	*Aten/pro halaxten* You went
Third	Masculine	*Hu/*pro halax* He went	*Hem/*pro halxu* They went
	Feminine	*Hi/*pro halxa* She went	*Hen/*pro halxu* They went

In normative Hebrew, the same paradigm applies for the future. In colloquial Hebrew, first person singular is nondistinct from third person singular, and thus, *pro* is licensed in the future only in first person plural, second person singular and second person plural, as shown in Table 4 for the verb *kataf* 'to pick up':

Table 4. The Paradigm for pro in Colloquial Hebrew Future Tense

Person	*Number* / *Gender*	*Singular*	*Plural*
First		*Ani/*pro yiktof* I will pick up	*Anaxnu/pro niktof* We will pick up
Second	Masculine	*Ata/pro tiktof* You will pick up	*Atem/pro tiktefu* You will pick up
	Feminine	*At/pro tiktefi* You will pick up	
Third	Masculine	*Hu/*pro yiktof* He will pick up	*Hem/*pro yiktefu* They will pick up
	Feminine	*Hi/*pro tiktof* She will pick up	

Thus, acquiring the use of null subject in Hebrew requires, in addition to the understanding of the term subject and its possible manifestation in the language (null or overt), a strong grasp of the morphological system.

Once children start combining words, the notion of subject becomes relevant. Armon-Lotem (1996) reports that the earliest two word combinations that contain a verb also contain a lexical subject, i.e., bare nouns and proper nouns, as in (8):

(8) *gaga halax* 'duck went' [Smadar 1;6]
doda Ogi yedafdef 'Aunt Orly will turn the page' [Leor 1;9]

At the same time they also use utterances without a subject, in a context based manner, reminiscent of topic-drop (Ariel 1990; Hyams and Wexler 1993). Gradually other types of subjects are added. On the one hand, children start using nominal sentences with a deictic *ze* 'it, this', and on the other hand, sporadic use of pronoun also become evident.

Berman (1990) discusses four phases in the use of overt subjects in Hebrew, using longitudinal data from three Hebrew speaking children. Two of these children have also been studied for the present paper, making it possible to track the relation between verbal morphology and the use of subjects, discussed in Section 4.3. The findings presented here, however, come from Berman's paper.

Berman reports that Hebrew speaking children initially use subjectless utterances with uninflected or unclear verbal forms, and once they start using overt subjects, they use bare nouns and proper names as in (9a), as well as a deictic ze 'this', as in (9b):

(9) a. *xatul nafal* 'cat fell', *Hagar (te)saper* 'Hagar (will) tell' [Hagar 1;7]

b. *ze anan* 'it (is a) cloud' [Smadar 1;6]

Only later do they start using personal pronouns and more elaborated subjects. In utterances with inflected verbs, the children's use of pronoun subjects alternates with the production of sentences with null-subjects. Many of these subjectless utterances are still ungrammatical, since they make use of present tense verbs which do not allow null-subject in Hebrew, as shown in (10):

(10) a. *hi* *shota* *mic*
she drink-pres.fm.sg. juice
'She drinks juice'

b. **holex* *ba-na'alayim* *shel doda Ogi*
walk-pres.ms.sg in-the-shoes of aunt Orly
'I am walking in the shoes of Aunt Orly' [Leor 1;11].

(10a) is grammatical since there is an overt third person subject, whereas (10b) is ungrammatical since the first person subject has been ungrammatically omitted.

Finally, a month or two later, children achieve an adult level of productive use of pronouns and grammatical use of null subjects. This sequence of acquisition – no subjects > bare nouns/proper nouns > pronouns > target-like null-subjects – which Berman reports for two of the children in the present study, Smadar and Leor, has also been found for the third child, Lior. This makes it possible to discuss the interrelations between the acquisition of inflections and the use of subjects for these three children.

4.3. Interrelating subject use and verbal morphology

Subject use and the acquisition of verbal morphology are expected to be interrelated since the morphosyntactic features of the verb have to agree with the subject. It can be predicted that if the child does not use overt subjects showing that she distinguishes between male and female subjects, she cannot be expected to use gender agreement appropriately. Similarly, if the child does not distinguish syntactically between herself and others, using subject pronouns, she cannot use person morphology agreeing with these subjects

These predictions are confirmed both by the order in which verbal morphology interrelates with subject use and by two types of agreement

mismatches found in the children's production. Figure 6 shows how the acquisition of verbal morphology interrelates with the subject use in the language production of one of the girls, Smadar:

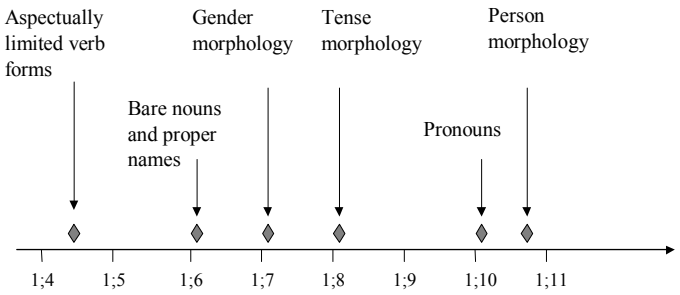


Figure 6. Interrelating subject use and verbal morphology (Smadar)

Figure 6 shows that Smadar uses aspectually limited verb forms from her first recordings, when she is 1;4. She starts using bare nouns and proper names at 1;6. A month later, when she is 1;7, she makes productive use of gender morphology, followed by upsurge in the use of tensed verbs at 1;8. Then, when she is 1;10 she uses pronouns, and finally, towards the end of this month she shows productive use of person morphology. Leor, the boy, makes productive use of tense when he is 1;11, but uses person inflections productively only when he is 2;1. Nevertheless, he uses pronouns ten days after he uses tense. This order of acquisition also holds for Lior, as well as for other children studied by Armon-Lotem (1996a).

This pattern in the acquisition of Hebrew, where bare subjects are used after aspect, but before gender and number agreement, and pronouns are used after tense but before person agreement, yields agreement mismatches, as in (11):

(11) a. **Hagar* *(na)fal*
Hagar(fm) fall-past.ms.sg
'Hagar fell' [Hagar 1;7]

b. **ani* *halax*
I-1st go-past.3rd
'I went' [Asaf 1;11]

(11a) is among Hagar's earliest word combinations when she is 1;7. At this point she makes aspectual distinctions, but has no gender agreement.

(11b) is used by Asaf, a boy, aged 1;11 from the Ravid corpus in CHILDES, when he is already using tenses, but has no person agreement.

Thus, both Hagar and Asaf seem to go through a stage where agreement morphology is missing, while aspect, and at a later stage tense, is manifested. Though these agreement mismatches are never found in more than one session per child, their consistent existence across the children in the database calls for an explanation.

These findings clearly indicate the central role of the finite verb, showing that the acquisition of the finite verb is related to the morphosyntactic features of the subject of the sentence. Though confirming the predictions for a relationship between subject use and verbal agreement, these findings raise some more questions. Why are bare subjects used after aspect, but before gender and number agreement? Why are pronouns used after tense but before person agreement? How are pronouns linked to tense? And what causes the agreement mismatches?

5. DISCUSSION – THE "MINIMALIST" CHILD

In order to answer the questions raised by the findings, we start by recalling the phrase marker proposed for Hebrew (Shlonsky 1995). The order of acquisition of verbal morphology suggests a bottom-up acquisition of this phrase marker, as shown in Figure 7:

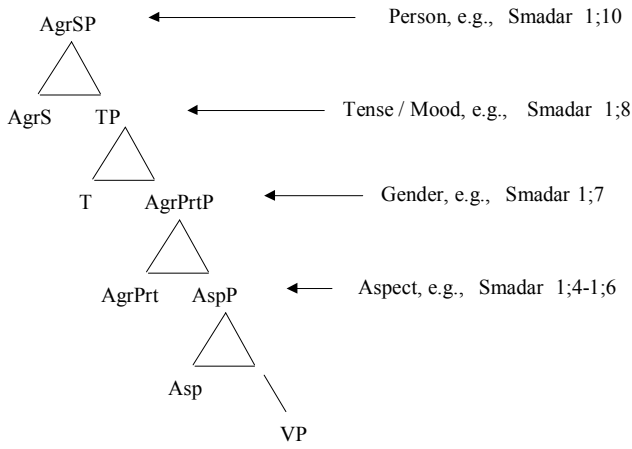


Figure 7. Bottom-up use of the functional system

The children use minimalist bottom-up trees, starting with the smallest subset of possible trees, a VP or a low aspectual phrase (Borer 1998), and gradually building up the tree, using AgrPrt for gender and number, then TNS, and only later AgrS for person agreement.

Assuming that children are minimalists, who acquire their morphology in a way which reflects the structure of the Hebrew phrase-marker, can explain both questions raised by the order in which inflectional morphology is acquired. Children start with aspect even when it is not marked syntactically in the input (the adult E-language), because aspect is universally the lowest available head, and in a language which does not mark syntactic aspect, this node can only be used for distinguishing telic from atelic actions (Borer 1998). Once aspectual distinctions are made using Asp, the inflections are gradually used as more functional heads are specified with features. This pattern shows a timing gap between the use of semantically motivated heads, such as Aspect or Tense, and their specifiers, the bare nouns and the pronouns, respectively. There is also a timing gap between the use of these specifiers and the use of syntactically motivated agreement heads. But a question remains as to which bootstrapping strategy is used in order to specify the head and spec features of these functional heads, and why there is such a crucial dependency between the different heads.

To this question are added a few more problems which are raised by the way verbal agreement and subject use are interrelated. The order in which subjects alternate with verbal morphology, that is, the use of bare nouns between aspect and gender agreement, and the use of pronouns between tense and person agreement cannot be accounted for without further stipulation on the nature of the Hebrew functional system. Nor does it explain the connection between the use of tense and the use of pronouns or the agreement mismatches.

One possible explanation is that while semantically motivated heads are head oriented, that is, the head is obligatory and its use is triggered via semantic bootstrapping, syntactically motivated heads are specifier oriented triggered via morpho-syntactic bootstrapping, since their use is motivated by the need to capture spec-head relations. Under such an account, the timing gap in the use of semantically motivated heads and the corresponding subjects can be attributed to the gradual use of the full range of functional categories, as suggested by the phrase marker in Figure 7. The agreement mismatches could be explained as an artefact of the specifier orientation of syntactically motivated functional heads, since the features of these heads can be specified only when the head is available, and the head in these cases is motivated by the specifier. This explanation is short, however, of capturing the correspondence between Asp and bare NP subjects and between TNS and

pronouns, since it is the specifier that motivates the head and not the head that limits the specifier.

In order to handle these issues, two "minimalist" working assumptions are used. First and foremost is the economy driven "no agreement nodes hypothesis" (Speas 1994; Chomsky 1994). Following Speas' (1994) proposal that the inventory of functional heads should ideally be limited to semantically motivated heads, and Chomsky's (1994) proposal that there should be no agreement nodes, since spec-head agreement is a purely syntactic notion, we would assume that only Tense and Aspect are valid functional categories at the IP level. The second assumption is that by locality principles, heads (and complements) are obligatory forming the most local domain, while specifiers are optional being in the local, but not most local, domain (Chomsky 1995). This assumption is necessary in order to explain the delayed use of specifiers (see also Clahsen, Eisenbeiss and Vainikka 1994).

In addition, a third assumption for acquisition is employed. It is assumed that children, being "minimalists" construct the smallest convergent trees that their grammar requires (Weissenborn 1993). These trees, as we saw, are expended in a bottom-up manner (Armon-Lotem 1996a).

Assuming that there are no Agr nodes, agreement features should be associated with the semantically motivated functional heads Asp and T. The crosslinguistic correlation between participial forms and aspect, suggests that Asp is specified with gender and number spec-features for participial agreement. Similarly, the correlation between tensed forms and person agreement, both of which are used deictically, in time and space, respectively, suggests that T is specified with person spec-features for subject agreement.

By the locality principles, heads are obligatory, specifiers (i.e. the external argument) are optional, and agreement is the mere syntactic relation between the two. The children are using the head before the specifier while agreement follows the use of both the head and the specifier, as is predicted by the locality principles. The head, being obligatory, is acquired via semantic bootstrapping and is used first with its complement (see Clahsen et al. 1994). Once the head is available to the child, syntactic bootstrapping based on positive evidence, e.g., the use of an initial overt subject before this head, trigger the use of the optional specifier. This extends the minimal tree from the most local domain to the local domain. Once both the head and the specifier are available, checking requires agreement between the two, evidence for which can be found via morphological bootstrapping.

In other words, by locality principles, children seem to be using the most local domain (Figure 8a) as soon as they start using multi-words utterances (see also Powers 1997):

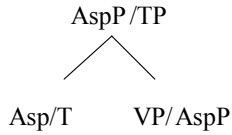


Figure 8a. Most Local Domain

Then, they extend their trees to the local domain where they have a subject (Figure 8b):

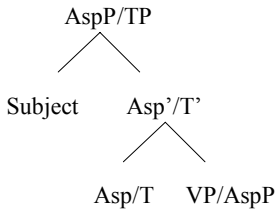


Figure 8b. Local Domain

And, finally, they define the syntactic agreement relations over this local domain, getting the desired agreement patterns which are typical of their language (Figure 8c):

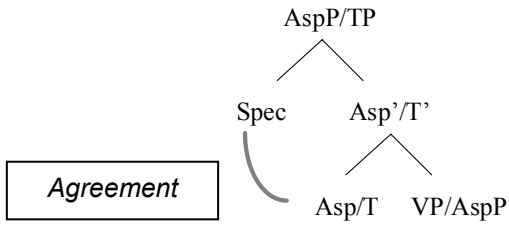


Figure 8c. Spec-Head Agreement

As shown in (8a-c), children initially use the semantically motivated head Asp. Next they project the optional specifier as a landing site for the external argument, the subject. Finally they apply the syntactic operation of checking spec-head agreement resulting in agreement in gender and number.

Using the "no Agr nodes" approach, the choice of the overt subjects associated with these heads suggests that due to checking the lack of a potential person feature on Asp prevents children from using a pronominal subject which is marked for person in the specifier of AspP. This can also

explain why a deictic *ze* 'this/it' which is not specified for person, is a possible specifier well before other pronouns show up. Only later, once T with its potential person feature is used by the child, can she use pronouns which are marked for person. When the semantically motivated functional head T is acquired, spec-head agreement in person applies in the same way. First, children project Tense which is semantically motivated as an anchor for deictic use of language in time. Then they project the optional specifier using pronouns for spatial deictic use. Finally they apply the syntactic operation of checking, resulting in agreement in person.

The agreement mismatches suggest that children figure out that Asp and T have a strong d-feature (Chomsky 1995) triggering movement very early on, prior to the appearance of agreement morphology on the verb.

6. CONCLUSIONS – THE CENTRAL ROLE OF VERBS

This paper shows that the finite verb has a central role in the acquisition of the phrase-marker, determining the features of the functional heads and the properties of the subject at each point in early syntactic acquisition. It shows that the interaction between the early use of verbal morphology and the early use of subjects can be accounted for by the "no Agr nodes" approach. Assuming that there are no Agr nodes, Asp is specified with gender and number spec-features and T is specified with person spec-features. Children start with the Aspect head, next they use the appropriate specifier, and then they apply the syntactic operation of spec-head agreement resulting in agreement in gender and number. When the semantically motivated functional head T is acquired, pronouns emerge in its specifier, and spec-head agreement in person applies. In this process children rely on semantic bootstrapping for identifying the semantically motivated heads, on syntactic bootstrapping for tracking down the specifiers, and morphological bootstrapping in order to identify the agreement relations which are manifested in their language.

In this manner of acquisition, children adhere to the same constraints on tree formation that adults do, as is predicted by the strong continuity hypothesis. Being more conservative and lacking the morpho-lexical knowledge of the particular paradigms of their language, children generate smaller convergent trees. With the growth in their morphological knowledge, children extend their use of functional categories, gradually specifying them for more features. This extension is always from semantically motivated head-features to syntactically motivated spec-features.

Thus, the relative order between the early use of verbal morphology and the early use of subjects gives further evidence for the central role of the finite verb in acquiring the morpho-syntactic features of constituents which

depend on the verb, e.g., the subject. It also sets a possible outline of the substages of the last stage of verb paradigms (Ingram et al. this volume). On the whole, this paper shows that "capturing the micro-steps in acquisition brings us a step closer to the articulation of a theory of UG that is biologically plausible" (Roeper 1994:30).

SHARON ARMON-LOTEM
Bar-Ilan University, Ramat Gan, Israel

7. NOTES

* This paper incorporates part of Armon-Lotem (1996a) and elaborates on earlier version of Armon-Lotem (1999). Research was supported by the Humboldt Foundation through an award to Ruth Berman (Tel-Aviv University) and Juergen Weissenborn (University of Potsdam). The Hebrew longitudinal data used in this paper were collected by the Tel-Aviv University Language Acquisition Project, as part of a crosslinguistic study of early word order directed by Ruth Berman and Juergen Weissenborn, supported by German-Israel Foundation grant no. I-11-070.4/87, by fundings from the DFG, and from the Child Language Data Exchange System at Carnegie-Mellon University, and from the Max Planck Institute for Psycholinguistics in Nijmegen, Holland.

[1] Both AgrPrt and AgrS check agreement with the subject. Historically, it was done only by AgrS, but in the model presented here the burden is shared by the two nodes.

8. REFERENCES

Ariel, Mira. *Accessing NP Antecedents*. London: Routledge, 1990.

Armon-Lotem, Sharon. The minimalist child: Parameters and functional heads in the acquisition of Hebrew. Doctoral dissertation, Tel-Aviv University, Israel, 1996a.

Armon-Lotem, Sharon. "What early Hebrew verbs teach us about root infinitives." In Charlotte Koster and Frank Wijnen (eds.) Proceedings *of 1995 GALA Conference*, University of Groningen, 1996b: 56-67.

Armon-Lotem, Sharon. "Agreement Mismatches and the Economy of Derivation." In Bart Hollebrandse (ed.), *New perspectives on Language Acquisition. University of Massachusetts Working Paper* 22. University of Massachusetts, Amherst, 1999: 27-36. no abbreviation

Armon-Lotem, Sharon and Ruth A. Berman. "The Emergence of Grammar: Early Verbs and Beyond." *Journal of Child Language* 30 (2003): 845-877.

Berman, Ruth A. *Modern Hebrew structure*. Tel Aviv: Universities Publishing, 1978.

Berman, Ruth A. "The Acquisition of Hebrew." In Dan I. Slobin (ed.), *The Crosslinguistic Study of Language Acquisition*. Hillsdale, NJ: Erlbaum, 1985: 255-371.

Berman, Ruth A. "A step-by-step model of language learning." In Iris Levin (ed.), *Stage and structure: Re-opening the debate.* Norwood, NJ: Ablex, 1986: 191-219.

Berman, Ruth A. "Acquiring an (S)VO language: Subjectless sentences in children's Hebrew." *Linguistics* 28 (1990): 1135-1166.

Berman, Ruth A. and Sharon Armon-Lotem. "How grammatical are early verbs?" In Claire Martinot (ed.), *Actes du Colloqu International sur l'Acquisition de la Syntaxe en Langue Maternelle et en Langue Etrangere. Annales Litteraires de l'Universite de Franche-Comte, 631* (1997): 17-59.

Bolozky, Shmuel. "The new imperative in colloquial Hebrew." *Hebrew Annual Review* 3 (1979): 17-24.

Borer, Hagit. *Parametric Syntax*. Dordrecht: Foris, 1984.

Borer, Hagit. "The morphology-syntax interface." In Andrew Spencer and Arnold Zwicky (eds.), *Morphology*. London: Basil Blackwell, 1998: 151-190

Chomsky, Noam. *The Minimalist Program*. Cambridge, MA: MIT Press, 1995.

Chomsky, Noam. *Bare phrase structure.* MIT Occasional Papers in Linguistics 5. Department of Linguistics and Philosophy, MIT, 1994.

Clahsen, Harald, Sonya Eisenbeiss, and Anne Vainikka. "The seeds of structure: A syntactic analysis of the acquisition of case marking." In Tuen Hoekstra and Bonnie Schwartz (eds.), *Language Acquisition Studies in Generative Grammar*. Amsterdam: John Benjamins, 1994: 57-84.

Grimshaw, Jane. Extended projection. Ms., Brandeis University, 1991.

Hyams, Nina and Ken Wexler. "On the grammatical basis of null subjects in child Language." *Linguistic Inquiry* 24 (1993): 421-459

Kayne, Richard. *The Antisymmetry of Syntax*. Cambridge, MA: MIT Press, 1994.

Powers, Susan M. The Growth of the Phrase Marker: Evidence from Subjects. Doctoral dissertation, University of Maryland, College Park, 1996.

Roeper, Tom. "Merge theory in acquisition." In Harald Clahsen (ed.), *Proceedings of Case and Agreement Conference*, University of Essex, Essex, 1994: 81-98.

Shlonsky, Ur. "Clause structure and word order in Hebrew: An essay in comparative Semitic syntax". Ms., University of Geneva, 1995.

Siloni, Tal. Noun Phrases and Nominalization. Doctoral dissertation, University of Geneva, 1994.

Speas, Margaret. "Null arguments in a theory of economy of projection." In Ellena Benedicto & Jeff Runner (eds.), *Functional Projections. University of Massachusetts Occasional Papers in Linguistics* 17, 1994: 123-151.

Vainikka, Anne and Marta Young-Scholten. "Gradual development of L2 phrase structure." *Second Language Research* 12 (1996): 7-39.

Weissenborn, Juergen. "Mommy's sock almost fits: Constraints on children's grammars." Paper presented at the 6th International Congress for the Study of Child Language, Trieste 1993.

Part 2

Language-specific variation in the development of predication and verb semantics

CHRISTINE CZINGLAR, ANTIGONE KATIČIĆ,
KATHARINA KÖHLER
AND CHRIS SCHANER-WOLLES

STRATEGIES IN THE L1-ACQUISITION OF PREDICATION: THE COPULA CONSTRUCTION IN GERMAN AND CROATIAN[1]

Abstract. Languages differ with respect to the appearance of a copula in (primary) predication structures: Languages like English, German and Croatian require the copula to be always overt, other languages like Hebrew and Russian do not express the copula in present tense contexts. Our comparative discussion of the acquisition of the copular construction in German and Croatian includes any occurrence of the verb *to be* (German *sein*, Croatian *biti*) together with a non-verbal main predicate. The copula is a purely grammatical device, expressing tense and agreement features. Like other grammatical categories it is often omitted in early child language; cf. Bloom (1970); Brown (1973); Radford (1990). In her study on the acquisition of the English copula, Becker (2000a, 2000b) suggests that copula omission is related to the so-called root infinitive phenomenon in English. Furthermore, she finds a connection between copula omission and the kind of predicate involved. We present new data from two German children (from 2;3 to 2;9) showing that German shows a very similar pattern of copula omission as English. But we find some differences between copulaless predication and the target-like copular construction that cannot be explained by simply adapting Becker's analysis for German. Just like English, German is classified as a root infinitive language; cf. e.g., Poeppel and Wexler (1993). In contrast to English and German, root infinitives have not been observed for early Croatian (cf. e.g., Katičić 1997, 2000). Our discussion of new data from one Croatian child (1;7 to 2;1) completes the picture: Copula omission does not occur at all. The hypothesized connection between root infinitives and copula omission is corroborated.

1. INTRODUCTION

In this paper, we present a comparative view on the first language acquisition of the copular construction in English, German and Croatian. Our discussion of new data from German and Croatian builds on the results of Misha Becker's study on the acquisition of the English copula (cf. Becker 2000a, 2000b). Note that by 'copular construction' we exclusively refer to occurrences of *to be* (German *sein*, Croatian *biti*) taking a non-verbal predicate, e.g., an adjectival phrase, a nominal phrase, or a prepositional phrase. The focus of our investigation is on the children's syntax and

N. Gagarina and I. Gülzow (eds.), The Acquisition of Verbs and their Grammar: The Effect of Particular Languages, 71–104.

morphosyntax of copular constructions, in contrast to other papers in this volume that treat comparable phenomena concentrating on the emergence and frequency of verb forms, cf. e.g., the contributions by Gülzow & Gagarina and Vihman & Vija.

When examining the development of the syntax and morpho-syntax of verbs, the copula is a particularly good case in point: It is a purely grammatical phenomenon, completely devoid of descriptive content. Nevertheless it occurs quite early and frequently in young children's speech across different languages as diverse as English (see Becker 2000a, 2000b) and Estonian (see Vihman & Vija, this volume). The copula plays a prominent role in the development of the mechanism of predication, i.e. the linking of a subject with its predicate. In the case of verbal predicates, both the syntactic and semantic part of this linking is accomplished by the finite lexical verb itself. If the main predicate of the sentence is not verbal, the copula takes over the syntactic part of predication: Its basic function is to carry tense and agreement features, i.e. to spell out finiteness, when it cannot be expressed otherwise.

During the last 20 years, theory building in language acquisition has centered around children's omission of functional material. It is a well established fact about child English that the copula is often omitted, resulting in sentences like *Mommy busy*; cf. e.g., Bloom (1970); Brown (1973), Radford (1990), Becker (2000a, 2000b). From a typological point of view, this is not just a child phenomenon. There are languages, for example Hebrew or Russian, that allow or even force the present tense copula to remain unexpressed.

In her extensive study on the acquisition of the English copula, Becker (2000a) argues that copulaless predication as a fact of first language acquisition is related to another pertinent characteristic of early child language, the so-called root infinitive phenomenon; see section 2.2.1 for more detailed information on root infinitives in the languages under discussion. Becker's conjecture is based on the old observation that the copula is almost always used in its finite form. Lexical verbs, on the other hand, alternate between non-finite and finite forms in the speech of young children acquiring English. Becker takes copulaless predication structures to be the non-finite counterpart of predicative structures involving an overt copula.

We investigate the acquisition of the copula in two other languages to test Becker's conjecture: German and Croatian. Like English, none of the two languages allows the copula to be missing in primary predication structures. Like English, German is a root infinitive language; cf. Poeppel and Wexler (1993); Schaner-Wolles (1994, 1995/96, 2000); Köhler and Bruyère (1995/96); Köhler (1998). Hence a similar pattern with respect to copula

omission is expected under the assumption that the two phenomena are related. Unlike English and German, the root infinitive phenomenon is not attested in Croatian; cf. Katičić (1997, 2000); Anđel et al. (2000). If the two phenomena are related, we expect predicative structures without an overt copula not to be attested either. Both expectations are borne out.

Becker's other important observation is that copula omission is not random.[2] There is a connection between the omission of the copula and the kind of predicate in English. We extend her query to the following more general questions: What are the differences between copulaless predication structures and those with a copula? Are those differences found across languages?

The article is organized in the following way: Section 2 contains theoretical considerations on predication and the copular construction, some observations on the acquisition of verb grammar in German and Croatian, a summary of Becker's study on the English copula and an outline of our research questions. In section 3, we present and compare longitudinal data from two children (age range from 2;3 to 2;9) acquiring Austrian German as their first language. The German data are compared to the results of Becker's study. Additionally, we review the properties of subjects and predicates in detail. Specific attention is devoted to the placement of subjects and predicates, respectively. In section 4, we present longitudinal data from one child (1;7 to 2;1) acquiring Croatian as a first language. We compare the data to the English and German findings.

2. BACKGROUND

2.1. The Structure of Predication

The linking of a subject with its predicate is a fundamental mechanism of language. With lexical verbs, it involves a semantic part (thematic roles) and a functional part (tense and agreement). The first part is usually associated with the lexical projection of the verb (VP). The second part deals with the functional projection(s) of the verb (IP or corresponding projections).

It is commonly assumed that the copula does not have semantic content, i.e. it does not assign thematic roles; cf. Déchaine (1993); Rothstein (1987); Williams (1980) and others. The semantic part of the predication must be accomplished by the non-verbal predicate. From a theoretical point of view, there are (at least) two options; cf. contributions in Cardinaletti and Guasti (1995). One possibility is to collapse the semantic and the functional part in their relation to the functional category of Infl(ection), i.e. the head of IP. The copula is treated on a par with the future auxiliary *will*, and the copular

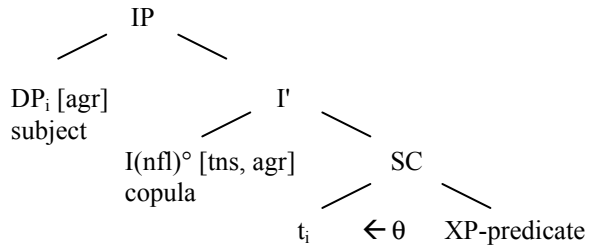
construction is seen as a simple instance of primary predication, the only difference being that XP is non-verbal. The following examples are from Williams (1984:136); see also Williams (1980), Déchaine (1993):

(1) a. John – will – leave. NP – Aux – VP
b. John – is – a fool. NP – Aux – NP

The second option is to separate the semantic part from the functional part. This option is referred to as the 'Small Clause Analysis'. Stowell (1978, 1983) assumes that the subject and the predicate form a lexical constituent, a so-called small clause (SC). In the SC-configuration the predicate assigns a thematic role to the subject. A Small Clause Analysis is also assumed for secondary predication constructions: The verb *consider* in (2a) takes a SC as its complement. In contrast to *consider*, the copula in (2b) does not assign its own thematic subject role, similar to raising verbs like *seem*. As depicted in (3) the subject of the small clause moves to the subject position of the 'big' clause (SpecIP), where it enters an agreement relation with the finite copula.[3]

(2) a. We consider [$_{SC}$ John clever]
b. John$_i$ is [$_{SC}$ t$_i$ clever]

(3)



In the following, we will restrict our discussion to non-verbal primary predication. The acquisitional data we present in this paper are in principle compatible with both analyses of copular constructions discussed. To be explicit, we will assume the Small Clause Analysis to be the underlying structure of the copular construction.

Languages differ in whether they allow primary predication without an overt copula or not. English and German, for example, always demand an overt copula. In Russian and Hebrew, on the other hand, an overt copula is not expressed in the present tense. It must be an option of universal grammar that finite main clauses can have an independent temporal interpretation without an overt verbal element.

None of the target languages we discuss allows the copula to be missing.[4] But English children produce predication structures without an overt copula that are not acceptable for adults. Radford (1990) argues that children's utterances like *Mommy busy* can be analyzed in terms of the 'Small Clause Hypothesis': Adopting Stowell's (1983) SC-Analysis, he assumes that the initial structures children produce are small clauses lacking functional structure. This position has been refuted by many researchers, who provide evidence for functional projections in early child language on the basis of data from other grammatical phenomena.

2.2. First Language Acquisition

The investigation of copula predication structures is situated within the larger area of the acquisition of verb grammar. To be explicit, by verb grammar we refer to the morphological properties of the verb as well as to its position with respect to the other elements in a sentence. Traditionally, verbs are viewed as the grammatical center of a sentence. In the field of generative syntax, the verb is a core case for the interplay between lexical and functional structure in the clause. Verbs are lexical heads, with lexical meaning and selection properties. They are inflected for morphosyntactic features, which are hosted by functional heads, a process that is reflected by verb movement in many languages. Verbs, their inflection, and their placement in early speech have attracted particular interest in the generative acquisition literature.

2.2.1. Verb Grammar in the Acquisition of German and Croatian

The present investigation of the copula construction in first language acquisition compares new data from German and Croatian with previous findings from the acquisition of English. German and Croatian differ in the acquisition of verb grammar. German children go through a stage that is characterized by the usage of verbs in the infinitive form, so called root infinitives.[5] The term "root infinitive" refers to contexts where finite main clauses are expected. During the same time span children also produce root sentences with target-like finite verb forms, cf. e.g., Poeppel and Wexler (1993); Schaner-Wolles (1994, 1995/96, 2000); Köhler and Bruyère (1995/96); Köhler (1998). The utterances in (4) are examples of root infinitives, the ones in (5) of finite clauses from the speech of the German speaking children investigated here:

(4)
a. Kind Kirschn essn (Nico 2;3)
child cherries eat-INF 'The child is eating cherries'
b. Mama Zug fahrn (Nico 2;4)
mama train ride-INF 'Mama is to ride the train'
c. Rutsche raufklettern (Paul 2;1)
slide up-climb-INF 'I want to climb the slide'
d. Pauli durchkriechn (Paul 2;4)
Pauli through-crawl-INF 'Paul wants to crawl through'

(5)
a. Hase kocht (Nico 2;3)
hare cook-3S 'The hare is cooking (sth.)'
b. Mama fahrt schon (Nico 2;4)
mama go-3S already 'Mama is already going'
c. wakal [: wagerl][6] faehrt (Paul 2;1)
cart.dim go-3S 'The little cart goes'
d. hamgetti [: spaghetti] macht Papa (Paul 2;4)
spaghetti make-3S Pap 'Papa makes spaghetti'

Croatian, on the other hand, is a language where the root infinitive phenomenon is absent from typical unimpaired development. This has been shown in several longitudinal studies; cf. Katičić (1997, 2000); Anđel et al. (2000). Furthermore, there are no parental reports about root infinitives. Consider the following examples of finite verb forms in very early speech:

(6)
a. Tu idem (Antonija 1;10.0)
here go-1S 'I am going here.'
b. Ja idem tamo (Antonija 1;10.0)
I go-1S there 'I am going there.'
c. Ti mene voliš? (Antonija 1;10.21)
you me.acc love-2S 'Do you love me?'

One observational difference between the acquisition of German and Croatian is that functional verbs (i.e. auxiliaries, modals, copula) come in earlier in Croatian than in German, where there is a considerable time lag between the appearance of lexical verbs and functional verbs. But the lexical-functional split plays yet another role in the acquisition of German. In root infinitive languages, the alternation between finite verb forms and the infinitive does not uniformly affect all verbs. During a time of linguistic development where children already use functional as well as lexical verbs, and still use root infinitives, they will usually display the finite form and infinitive alternation only with lexical verbs. The finite – infinitive alternation will appear extremely rarely or not at all with functional verbs, the generalization being that functional verbs appear in finitely inflected

forms only.[7] Besides this observation, there is presumably more to be said about which verbs do or do not appear as root infinitives.[8] But this topic is beyond the scope of this paper. Still, across different root infinitive languages and different children, the copula in particular is noted to be virtually absent from root infinitives.

2.2.2. Becker (2000a, 2000b): *A Study on the Copula in Child English*

Children acquiring English frequently omit the third person singular *–s,* the only inflectional morpheme in the regular present tense paradigm. They produce sentences like (7a), but finite utterances like (7b) and (7c) also appear at the same time. The examples in (7) are available in CHILDES, Brown (1973), examples a. and b. quoted in Guasti (2002):

(7)	a.	Cromer wear glasses	(Eve 2;0)
	b.	I don't want soup	(Eve 1;11)
	c.	My finger hurts	(Eve 2;0)

Although English does not have a morphological marking for the infinitive, (7a) arguably has a similar status as root infinitives in German. German infinitives bear overt morphological marking, which is homophonous with two finite suffixes (1P and 3P). Syntactically, the English and German infinitives found in children's utterances behave exactly like infinitives. So English is categorized as a root infinitive language; Wexler (1994). Vihman & Vija (this volume) discuss the acquisition of verb inflection in Estonian, a language that has stems and morphologically marked infinitives. They describe a stage in the acquisition process where stems are used instead of finite forms, but the question whether these are to be categorized as root infinitives is not discussed.

In English child language the copula almost always appears in its finite form. Starting from this observation, Becker (2000a, 2000b) dedicates an extensive study to the copula in child English. The investigation is based on data from five children available from the CHILDES database (cf. MacWhinney and Snow 1990). Like Bloom (1970), Brown (1973), Radford (1990) and others, Becker observes that copula predication sentences alternate with utterances that combine a subject and a predicate, but do not contain a copula. Examples (8) through (10) illustrate this: [9]

(8)	a.	You're so dirty	(Naomi 2;0)
	b.	me tired	(Naomi 2;0)

(9) a. Daddy's at school (Peter 2;2)
b. I in the kitchen (Nina 2;1)

(10) a. he's a dog (Nina 2;0)
b. I not honey. I Adam Smith (Adam 2;11)

Becker's central observation is that there is a connection between the omission of the copula and the kind of predicate involved. In particular, a contrast is found between nominal predicates in (10) and locative predicates as (9): Most nominal predicates appear in predication structures with a copula. On the average, the copula is present in 66% of the nominal predication cases (out of a total of 1,155 utterances). With locative predicates, copula omission is predominant: The copula is present in only 28% of the cases (out of a total of 295 utterances).

Becker (2000a, 2000b) argues that this is due to a semantic difference between nominal and locative predicates. Interactions of form and meaning in language acquisition are also invoked in Tsujimura (this volume) on motion verbs in Japanese. In Becker's analysis, semantic properties of predicates influence the way they are syntactically embedded in the clause. Nominals are core cases of temporally unbounded properties, or individual-level predicates, and locatives are core cases of temporally bounded properties, or stage-level predicates; cf. Carlson (1977, 1980) and Kratzer (1988, 1995). Becker (2000a, 2000b) assumes a syntactic analysis in which sentences with stage-level predicates contain additional functional structure: an Aspectual Phrase (AspP)[10] above the predicative phrase which hosts an event argument in the sense of Davidson (1967).[11] In accordance with a proposal by Kratzer (1988, 1995), Becker assumes that individual-level predicates lack such an event argument and therefore have no AspP. Following Enç (1987), she assumes that temporal anchoring of main clauses is accomplished by a Tense operator binding Infl in adult English. Becker argues that during the root infinitive stage this temporal anchoring requirement can be fulfilled by binding Asp, i.e. the head of AspP. In the case of non-verbal predication structures, the copula is the spellout of finite Infl, while Tense binding of Asp yields no such result.[12] The English learning child is forced to use the copula when Infl is the only head available for binding by the Tense operator, i.e. in the case of individual-level predicates. Stage-level predicates, on the other hand, offer another possible bindee, Asp. This explains the higher rate of copula omission with stage-level predicates and the lower rate of copula omission with individual-level predicates in child English.[13]

2.2.3. Research Questions

In the rest of this paper, we pursue the following research questions concerning the copula in the acquisition of German and Croatian: Given that children acquiring German use root infinitives, like children acquiring English – do they also use copulaless predication structures? As this is indeed the case, the second question ensues: Are there differences in frequency of copula omission for different kinds of predicates? We will show that this is so in German, too. German and Croatian contrast with respect to the root infinitive phenomenon: There are no root infinitives attested in the acquisition of Croatian. Becker assumes that copulaless predication structures are the root infinitive counterpart of the overt copula. This predicts that children acquiring Croatian should not use copulaless predication structures. We will show that this prediction is actually borne out. Furthermore we explore what other similarities and differences there are between predication structures with and without a copula.

3. GERMAN

3.1. Target Language

In this section, we give a brief overview of some core characteristics of German that are relevant to the understanding of the discussion on the acquisition of copula predication structures. Section 3.1.1 describes basic word order facts, section 3.1.2 informs about the copula predication structures in German.

3.1.1. Basic Clause Structure and Verb Placement

German main and embedded clauses differ in word order quite sharply. Declarative main clauses are often SVO, cf. (11a), embedded clauses are usually SOV, cf. (12). The verb second property of German further allows the main clause word order OVS like in (11b), among others. The analysis of German word order in the generative tradition assumes a common basic word order behind these surface patterns; the other word orders are considered to be derived from the basic word order by movement. At least since Koster (1975) it is commonly assumed that the embedded clause order SOV is the underlying word order. The main clause SVO order is derived by movement of the finite verb to second position. The position to the left of the finite verb has as special status that distinguishes verb second languages from non-verb second languages. Any kind of constituent may occupy this position, but never more than one. In the unmarked case, this constituent is the subject.

The subject appears to the right of the finite verb if some other constituent appears in first position, like e.g., in the OVS word order of (11b). The movement of a constituent to the clause initial position is referred to as 'topicalization':

(11) a. Maria liest das Buch
Mary reads the book 'Mary is reading the book'
b. Das Buch liest Maria
the book reads Mary 'Mary is reading the book'

(12) Hans glaubt, dass Maria das Buch liest
Hans thinks that Mary the book reads
'Hans thinks that Mary is reading the book'

3.1.2. The Copula

In German, the copula *sein* occurs with a variety of non-verbal predicates[14]. The examples in (13) are subordinated clauses representing the basic word order, the copula is in italics:

(13) a. *adjective*
dass Peter brav *ist*
that Peter good is
'that Peter is good'
b. *nominal*
dass ihr Bruder (ein) Lehrer *ist*
that her brother (a) teacher is
'that her brother is a teacher'
c. *locative*
dass die Kinder im Hof *sind*
that the children in-the yard are
'that the children are in the yard'
d. *particle*
dass Peter schon weg *ist*
that Peter already away is
'that Peter is already gone'

'Adjective', 'nominal', and 'particle' refer to lexical categories, whereas 'locative' is a semantic category, which comprises locative PPs just as well as adverbs.

3.2. Data

Our counts for German were carried out on spontaneous production data from two longitudinal child language corpora: The Nico-Corpus, collected by Sabine Bruyère, cf. Köhler and Bruyère (1995/96), and the Paul-Corpus, collected by Katharina Korecky-Kröll, cf. Korecky-Kröll (2000). Nico and Paul are boys growing up in Vienna, Austria, acquiring the local variety of spoken German. The time span chosen for a thorough analysis was about five months for both children. This is the interval in which predication constructions with and without copula co-occur.

3.2.1. Analyzed Utterances

In order to be included in the analyses, utterances had to fulfill the following criteria: Only declarative multi-word utterances were included. Utterances with a copula were included if they consisted of at least a subject and a predicate. We define copulaless predication structures as combinations of a subject and a predicate that have the same meaning as this subject predicate combination would have with a copula, i.e. stative. Imitations and self-repetitions were excluded. This results in a sample of 700 utterances of Nico (age period: 2;3.30-2;9.4), and 108 utterances of Paul (age period: 2;4.6-2;8.21). For details see Appendix A.

3.2.2. Classification of Subjects versus Predicates

In copular sentences with two nominal or pronominal constituents, the two are potentially ambiguous with respect to subject or predicate status. In this case we adopt the following general heuristics (cf. Williams 1997; Löbel 2001). Predication is asymmetrical: The subject is more referential than the predicate, in the sense that the speaker is more directly acquainted with its referent. The predicate noun phrase specifies a property of this referent. As a matter of fact, many child utterances contain demonstrative pronouns with a clear situational (deictic) interpretation, which we consequently classified as subjects. For the few sentences with two bare noun candidates for the subject and predicate roles, we relied on the context information. In the following example, the context makes it clear that the car Nico is talking about functions as an aircraft, and not vice versa:

(14) auto flugzeug *is* (Nico 2;7.24)
car airplane is 'The car is an airplane'

3.3. Copular Constructions in Child German

3.3.1. Predication with or without Copula

Both children use predication structures with and without a copula. In Table 1, the proportions of simple predication structures without an overt copula are 46.9% in the Nico-Corpus and 36.1% in the Paul-Corpus.

Table 1. Overview of analyzed data from Nico and Paul, with overt subject (absolute numbers and row percentages in brackets)

		copula	*no copula*	*total*
Nico	(2;3.30 - 2;9.4)	372 (53.1%)	328 (46.9%)	700 (100%)
Paul	(2;4.6 - 2;8.21)	69 (63.9%)	39 (36.1%)	108 (100%)

See (15) for examples of predication with and without a copula (predicates are underlined).

(15) a. das *is* die Mama (Paul 2;7.15)
that is the mama 'That's mum.'
b. keksi kueche noch (Nico 2;6.12)
cookie kitchen still 'The cookie is still in the kitchen'

The generalization that the copula hardly ever occurs in the infinitive also holds in our sample: 369/372 or 99.2% finite copula forms were counted for Nico, and 69/69 or 100% finite forms for Paul.

For both children, the coexistence of copula and copulaless predication structures falls in the same period as the coexistence of finite verbs and root infinitives (cf. Köhler and Bruyère 1995/96; and Schaner-Wolles 2000 for Nico; Korecky-Kröll 2000 for Paul). Figures depicting the overall development of the finite verb and root infinitive ratios for Nico and Paul are given in Appendix B.

3.3.2. Predicates

Constituents classified as predicates were grouped into four categories: Adjectival, locative and nominal predicates are distinguished on semantic and categorical grounds, the fourth category ('other') is a combination of

constituents that could not be classified in the way described above (adverbials, particles, interjections, onomatopoetic expressions, child-specific words). Table 2 indicates that all kinds of predicates occur with and without an overt copula. While predication structures without a copula decrease steadily from Nico18 (2;8.7) and Paul11 (2;5.16) onwards, there is no development in the distribution of predicate types, the main topic to be discussed in this section. Therefore the data from all the recordings of each child have been pooled for quantitative analysis. In this sense, our study contrasts with other studies in this volume that concentrate on longitudinal development, e.g. Ingram, Welti & Priem; Gülzow & Gagarina; Vihman & Vija; Tsujimura. File by file information on the Nico and Paul data is given in Appendix C.

Table 2. Predicate types and occurrence of the copula (absolute numbers and total row percentages)

	Nico06-20 (2;3.30 - 2;9.4)			*Paul09-17 (2;4.6 - 2;8.21)*		
predicate	*cop.*	*no cop.*	*total*	*cop.*	*no cop.*	*total*
adjectival	52	46	98	13	8	21
locative	61	168	229	26	20	46
nominal	237	77	314	23	4	27
other	22	37	59	7	7	14
total	372	328	700	69	39	108
	53.1%	46.9%	100%	63.9%	36.1%	100%

The following are examples of adjectival (16), locative (17) and nominal (18) predicates with and without an overt copula. The category 'other' is exemplified by an adverbial in (19a) and an interjection in (19b).

(16) a. Picki *is* krank (Nico 2;7.24)
Picki is ill 'Nico is ill'
b. schmutzig duplokiste (Nico 2;6.12)
dirty duplo-box 'The duplobox is dirty'

(17) a. kuchn *is* drinnen (Nico 2;3.30)
cake is inside 'The cake is inside'
b. keksi kueche noch (Nico 2;6.12)
cookie kitchen still 'The c. is still in the kitchen'

(18) a. das *is* die Mama (Paul 2;7.15)
that is the mama 'That's mum.'
b. den da polizei (Nico 2;7.3)
the-SG:MASC:ACC there police 'This is the police there'

(19) a. jetzt *is* der da dran (Nico 2;7.3)
now is the-SG:MASC there ADV
'Now it is this one's turn'
b. sockn igitt (Nico 2;6.1)
socks yucky 'The socks are disgusting'

Whereas locative and nominal predicates have a strong correspondence to stage-level and individual-level semantics respectively, adjectives are not uniform with respect to this distinction. Some adjectives have an obvious semantics (e.g., *hungrig* 'hungry' is a stage-level, *blond* 'blonde' an individual-level adjective), and others permit both interpretations depending on the context (e.g., *kaputt* 'broken'). Furthermore, especially in the early files, adjectives are not very frequent in either corpus. Similar difficulties arise in Becker's work; cf. Becker (2000a:132ff). For these reasons, we will not take adjectives into consideration in this paper.

As far as locative and nominal predicates are concerned, our data confirm Becker's (2000a, 2000b) observation on English predicatives for German: Locative and nominal predicates differ with respect to the frequency of occurrence of an overt copula. In the Nico-Corpus, the majority of locative predicates (73.4%) occur without a copula, and the majority of nominal predicates (75.5%) occur with an overt copula. The distribution is not so clear for Paul's locatives, but it is even clearer for his nominal predicates. Still, the difference between the two predicate types is statistically significant according to the chi-square test.

Table 3. Locative vs. nominal predicates – occurrence of the copula (absolute numbers and row percentages)

	Nico06-20 (2;3.30 - 2;9.4)			*Paul09-17 (2;4.6 - 2;8.21)*		
predicate	*cop.*	*no cop.*	*total*	*cop.*	*no cop.*	*total*
locative	61	168	229	26	20	46
	26.6%	73.4%	100%	56.5%	43.5%	100%
nominal	237	77	314	23	4	27
	75.5%	24.5%	100%	85.2%	14.8%	100%
total	298	245	543	49	24	73
	54.9%	45.1%	100%	67.1%	32.9%	100%
	chi-square test sign.: p=.000			chi-square test sign.: p<.05		

Among the locative predicates, full fledged prepositional phrases are extremely rare (3 out of 229 locative predicates are PPs for Nico, 3 out of 46 for Paul). Locative adverbs constitute the majority of locative predicates, with *da* 'there' being the most frequent locative adverb. Paul uses 43 locative adverbs, Nico 158. There is a difference between the boys insofar as Nico also uses bare noun phrases with locative meaning. They appear in 68 of his 229 locative predication structures (67 of them are bare nouns, only one case contains a determiner). A typical example of a locative nominal is given in (17b) above. Note that the vast majority of Nico's locative nominals (59/68 or 86.8%) is found in utterances without a copula and only 9/68 appear together with an overt copula. It seems that non-target-like structures avoiding functional elements such as the copula and prepositions cluster together.

A similar observation holds for nominal predicates. Apart from 5 pronominal predicates in the Nico-Corpus, nominal predicates are descriptive lexical NPs.[15] Most lexical NP-predicates are bare nouns, but some also contain functional material, i.e. a determiner (70/309 or 22.7% for Nico, 10/27 or 37.0% for Paul). Again, we observe that non-target-like properties cluster together: Lexical NP-predicates in copulaless predication structures are much less likely to appear with a determiner than those in the target-like construction. Of 77 nominal predicates without a copula, Nico uses only 5 with a determiner (6.5%), whereas 65/232 or 28.0% of the nominal predicates which are found in the context of an overt copula contain a determiner. For Paul the numbers are too small for meaningful comparison (1/4 or 25.0% determined NP-predicates without a copula and 9/23 or 39.1% with a copula).

3.3.3. Subjects

Constituents classified as subjects were grouped into two categories: Pronominal subjects and descriptive lexical noun phrases (cf. endnote 15) including nominalized adjectives.

Both subject types appear in children's predication structures, whether they contain an overt copula or not. But when we look at pronominal subjects, we can see that they are strongly preferred in the context of the target-like predication structure with an overt copula (79.2% for Nico, 82.6% for Paul).[16]

Table 4. Types of subjects in contexts with overt copula and without copula (absolute numbers, row and column percentages)

	Nico06-20 (2;3.30 - 2;9.4)			*Paul09-17 (2;4.6 - 2;8.21)*		
subject	*cop.*	*no cop.*	*total*	*cop.*	*no cop.*	*total*
pronominal	312	82	394	38	8	46
row pct.	79.2%	20.8%	100.0%	82.6%	17.4%	100.0%
col. pct.	83.9%	25.0%	56.3%	55.0%	20.5%	42.6%
lexical	60	246	306	31	31	62
row pct.	19.6%	80.4%	100%	50.0%	50.0%	100.0%
col. pct.	16.1%	75.0%	43.7%	45.0%	79.5%	57.4%
total	372	328	700	69	39	108
row pct.	53.1%	46.9%	100.0%	63.9%	36.1%	100.0%
col. pct.	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%

Most pronominal subjects are demonstrative pronouns, i.e. weak demonstratives (*der/die/das* and the reduced form *(da)s*), the morphological paradigm of which is almost identical to the definite determiner. As allowed in the target language, weak demonstratives are sometimes reinforced by the locative adverb *da* 'there' (for example *das da* 'that there'). By far the most frequent subject pronoun is the singular neuter form of the weak demonstrative *das* 'this/that/the.n.sg' (293/394 or 74.4% for Nico, 35/46 or 76.1% for Paul); see (15), repeated here as (20):

(20) das *is* die Mama (Paul 2;7.15)
that is the mama 'That's mum'

As regards predication structures with an overt copula, we find that lexical and pronominal subjects are evenly distributed in Paul's data. Nico,

on the other hand, displays a striking asymmetry with respect to pronominal vs. lexical subjects (83.9% of his copular constructions contain a pronominal subject). To some extent, this reflects his preference for one particular word order pattern: *X is das* 'X is that'. In total, Nico produced 161 utterances of this form, where X stands for any kind of predicate (plus sometimes additional material). Paul does not share this strong preference (9 occurrences); see also section 3.3.4 below. As for copulaless predication structures, most of their subjects are descriptive lexical noun phrases (75.0% for Nico and 79.5% for Paul) as in (16b) above. Again, it seems that non-target-like structures avoid functional elements, also in the case of pronouns.

A similar observation holds for determiner placement in lexical subjects; cf. endnote 15. Most lexical subjects are bare nouns, but some also contain a determiner (42/306 or 13.7% for Nico, 19/62 or 30.6% for Paul). In copulaless predication structures lexical subjects are less likely to appear with a determiner than in the target-like construction. Of 246 lexical subjects without a copula, Nico uses only 15 with a determiner in copulaless structures (6.1%), whereas 27/60 or 45.0% contain a determiner in the context of an overt copula. Paul uses 5/31 or 16.1% determined lexical subjects without a copula and 14/31 or 45.2% with a copula.

3.3.4. Placement of Subjects and Predicates

In this section, we investigate the question whether the placement of subjects and predicates in children's predication structures can tell us anything about the underlying clause structure, in particular in contexts without a copula. As already mentioned, it has been suggested that children's utterances lacking functional material might reflect a generalized small clause strategy; cf. Radford (1990). We want to find out whether copulaless (i.e. verbless) predication structures provide evidence for the presence of functional structure above a small clause in this clause type.

Most predication sentences containing a copula are actually V2 sentences (339/372 or 91.1% for Nico, 61/69 or 88.4% for Paul). Because of V2 and the availability of topicalization, the order between subjects and predicates is not fixed in German main clauses. Any kind of predicate can move across the subject to SpecCP (cf. section 3.1.1. above). Therefore, children are expected to make use of the topicalization possibility in a root clause with the copula in second position. We know independently that both Nico and Paul have command of the V2 property and topicalize constituents to the preverbal position, cf. Köhler and Bruyère (1995/96); Korecky-Kröll (2000). So we have all the more reason to expect predicate topicalization in copula V2 clauses. And indeed, both children use predicate-initial clauses with all kinds

of predicates. Nico displays a peculiar preference for predicate topicalization over subject-initial orders (in 62.8% of the cases). Paul, on the other hand, does not seem to prefer one over the other (predicate-initial orders appear in 45.9% of the cases). As mentioned above, Paul does not use the particular word order pattern that Nico uses very frequently, *X is das* 'X is that', as in (21):

(21) autobus *is* das (Nico 2;3.30)
bus is that 'This is a bus'

As this might blur the comparison between the two children, we separated the pattern *X is das* in Table 5. In that way, the similarity of the topicalization rates for both children becomes evident.

Table 5. Copula V2 clauses: distribution of subjects, predicates, and other constituents in initial position
(absolute numbers and column percentages, including and excluding 'X is das')

	Nico06-20 (2;3.30 - 2;9.4)			*Paul09-17 (2;4.6 - 2;8.21)*		
initial constituent	*#*	*%*	*% exc. 'X is das'*	*#*	*%*	*% exc. 'X is das'*
subject	123	36.3	69.1	31	50.8	59.6
predicate:						
'X is das'	161	47.5	–	9	14.8	–
other predicate	52	15.3	29.2	19	31.1	36.5
other constituent	3	0.9	1.7	2	3.3	3.8
total V2	339	100.0		61	100.0	
total exc. 'X is das'	178		100.0	52		100.0

(22a) and (22b) are examples of subject-initial and predicate-initial orders in V2 clauses with an overt copula:

(22) a. das *is* die Mama (Paul 2;7.15)
that is the mama 'That's mum'
b. da *is* eh lastauto. (Nico 2;5.11)
there is ADV truck 'There is a truck anyway'

We now turn to predication structures without a copula. Under the assumption that copulaless predication is represented as a bare lexical small clause with no functional projections (cf. Radford 1990), predicate-initial

orders are not expected. Table 6, however, shows quite a number of predicate-initial orders, contrary to this prediction.

Table 6. Relative order of subject and predicate in predicative utterances without copula, all predicate types (absolute numbers and row percentages)

Nico06-20 (2;3.30 - 2;9.4)			*Paul09-17 (2;4.6 - 2;8.21)*		
su (...) pr	*pr (...) su*	*total*	*su (...) pr*	*pr (...) su*	*total*
248	80	328	27	12	39
75.6%	24.4%	100.0%	69.2%	30.8%	100.0%

(23a) and (23b) are examples of subject-initial and predicate-initial orders in copulaless predication structures:

(23) a. den da polizei (Nico 2;7.3)
the-SG:MASC:ACC there police 'This is the police there'
b. fertig Picki (Nico 2;4.10)
ready Nico 'Nico is done'

Another conceivable hypothesis is that the clause structure of predication without a copula is just like that of an overt copula construction, only without a phonetic spellout of the feature bundle associated with the copula. Under this hypothesis the situation with respect to the relative order of subjects and predicates should be identical in both clause types. The evidence is equivocal, as the inversion rates in copulaless predication structures are below the topicalization rates in V2-sentences with a copula. But the difference is not big (24.4% compared to 29.2% for Nico, 30.8% compared to 36.5% for Paul). Also bear in mind that there are other differences between the two clause types that still call for an explanation.

In sum, we find that the comparison between predicative utterances with and without a copula with respect to subject and predicate placement is inconclusive. The question of how to account for children's predication structures without an overt copula is not settled yet.

3.3.5. Summary for German

We find the following systematic contrasts between predication structures with and without a copula in child German: When the copula is overt, it is almost always finite and mostly in the correct V2 position. Structures without a copula do not display a visible sign of finiteness, i.e. are classified as not

finite. As for the role of predicate types, Becker's observation is confirmed for German: Locative predicates prefer the structure without an overt copula, nominal predicates prefer the target-like copular construction. But there are more differences between predication structures with and without a copula. Non-target-like structures cluster together, so Nico's bare noun locatives strongly prefer predication contexts without an overt copula. Functional elements like (subject) pronouns or determiners are more likely to appear in predication structures with an overt copula than in its non-target-like copulaless counterpart. Although predicate inversion is available in both construction types, the predicate precedes the subject more frequently in the target-like construction containing a finite copula. We believe that these differences call for a structural explanation.

Table 7. Summary of the properties of predication structures with and without a copula in early child German

	structure with copula	*structure without copula*
copula		
finiteness	finite	not finite
placement	V2	(does not apply)
predicate		
locative predicates	dispreferred context	preferred context
nominal predicates	preferred context	dispreferred context
Nico's bare N locatives	dispreferred context	strongly preferred context
subject		
lexical subjects	infrequent	frequent
pronominal subjects	frequent	infrequent
DPs		
determiners (dets)	some dets occur	very few dets occur
word order		
predicate inversion	available/more frequent	available/less frequent

4. CROATIAN

4.1. Target Language

It is a crucial characteristic of Croatian that verbal inflection can identify sentential subjects which can therefore be null. In the variety acquired by the child overt pronominal subjects are not necessarily connected with an emphatic meaning (cf. Magner 1966:34). The following examples show the usage according to the standard language.

(24) a. idemo u grad
go-1P to town 'We are going to town'
b. mi idemo u grad
we go-1P to town 'WE are going to town'

The copula is realized by clitic, strong and negated forms of the verb *biti* 'to be' as shown in (25).[17] Strong affirmative forms have emphatic interpretation. Negated forms are always strong.

(25)

	biti 'to be'		*ne biti 'not to be'*
	clitic	*strong*	*negated*
1.sg.	sam	jesam	nisam
2.sg.	si	jesi	nisi
3.sg.	je	je[18]	nije
1.pl.	smo	jesmo	nismo
2.pl.	ste	jeste	niste
3.pl.	su	jesu	nisu

Although Croatian is considered to have free word order, the clitic copula is fixed to the second position of the clause. When it appears together with other 'second position' clitics (e.g. pronouns), they form a clitic cluster in which they are subject to ordering restrictions.[19]

The copula appears with various types of non-verbal predicates (cf. Barić et al. 1997:401-403). Following the categorization for German predicates we focus on adjectives, nominals and locatives as predicates; cf. (26). These are also the predication types used most frequently by the child. Moreover, Croatian adjectival predicates show number and gender agreement with their subjects; cf. (26a) and (26a').

(26) a. *adjective*
Jelo je vruće
food-NEUT is-CLIT hot-NEUT 'The food is hot'
a'. *adjective*
Juha je vruća
soup-FEM is-CLIT hot-FEM 'The soup is hot'
b. *nominal*
Merzedes je žena
M. is-CLIT woman 'Merzedes is a woman'

c. *locative*
To je u kući
that is-CLIT in house-LOC 'That's in the house'

4.2. Data

The Croatian data used for this study are part of a longitudinal corpus documenting the first language acquisition of a monolingual girl growing up in Zagreb. Antonija's spontaneous productions have been recorded for 20 minutes on an average and transcribed regularly, three to four times per month, by Draženka Blaži.

The analysis covers all sessions within seven months from the onset of the child's two-word stage (age period: 1;7.2 - 2;1.28). We included exclusively declarative sentences and also one word contexts. Repetitions, songs, nursery rhymes, imitations and interrupted utterances were excluded from the calculation. In total, 17 recordings and 97 of Antonija's utterances with a copula were analyzed; for details see Appendix D.

4.3. Copular Constructions in Child Croatian

Our description of the Croatian data is structured as follows: The core finding is that there are no cases of copula omission (section 4.3.1.). Therefore the sections 4.3.2 through 4.3.4 deal only with the properties of predicative structures with a copula.

4.3.1. No Omissions of the Copula

Already from the beginning of the two-element stage (1;7) Antonija produces copular constructions. The most striking observation compared to the findings from German and English is that we find no instances of non-verbal predication without a copula at all. Furthermore, we find that Antonija corrects herself when she leaves out the copula as it is shown in (27).

(27) to kajun – to je kajun [: klaun] (Antonija 1;9.15)
that clown – that is-CLIT clown 'That is a clown'

4.3.2. Utterances with a Copula

In general, we observe that Antonija's usage of the copula conforms to the regularities of the target language. This concerns primarily the contextual use of clitic, emphatic or negated forms and the placement of the clitic copula at

the second position of the clause. Furthermore, subject and verb agreement shows no deviations from the target language. Clitic clusters are very rare and they do not appear at all with copulas.

Table 8 shows the quantitative distribution of clitic and non-clitic (strong and negated) forms of the copula. The dominance of enclitic forms (almost 75% of all produced copulas) indicates that the child chooses enclitic forms for the neutral (non-emphatic) affirmative context.

Table 8. Clitic, non-clitic and negated forms of the copula, Antonija07-23 (1;7.2 - 2;1.28), (absolute numbers and row percentages)

clitic	*strong*	*negated*	*total*
72 (74.2%)	12 (12.4%)	13 (13.4%)	97 (100%)

Table 9 provides figures of all types of copular constructions taken into consideration. They include utterances with overt subjects and non-verbal predicates, null subject constructions with overt predicates, one utterance with an overt subject without an overt predicate and, finally, copulas as one-word utterances. The latter can be interpreted as elliptic comments or answers.

Table 9. Types of utterances with a copula, Antonija07-23 (1;7.2 - 2;1.28), (absolute numbers and row percentages)

cop, su, pr	*cop, pr*	*cop, su*	*cop*	*total*
48 (49.5%)	31 (32.0%)	1 (1.0%)	17 (17.5%)	97 (100%)

4.3.3. Predicates

Copular constructions with predicates are among the first two-word utterances. The earliest predicate types to appear are nominals and locatives, whereas adjectival predicates show up later, at 1;9. The most frequent locative predicates are adverbs represented in 21 cases by the deictic *tu* 'here'. Table 10 shows the distribution of the various predicate types.

Table 10. Predicate types, Antonija07-23 (1;7.2 - 2;1.28), (absolute numbers and row percentages)

adjectival	*locative*	*nominal*	*other*	*total*
19 (24.1%)	27 (34.2%)	31 (39.2%)	2 (2.5%)	79 (100%)

The following examples show adjectival (28), locative (29), nominal (30) and 'other' (31) types of predicates in the child.

(28) a. deda je umoran (Antonija 1;10.0)
grandpa is-CLIT tired-MASC:SG
'Grandpa is tired'
b. ti si jočeta [: zločesta] (Antonija 1;11.17)
you-SG are-SG:CLIT bad-FEM:SG
'You are bad' (said to grandma)

(29) a. tu je koka (Antonija 1;10.10)
there is-CLIT hen
'There is the hen'
b. u omaju [: ormaru] je (Antonija 1;9.15)
in cupboard-loc is-CLIT
'(the plate) is in the cupboard'

(30) a. i ja sam t(v)oja mama (Antonija 1;10.21)
and I am-CLIT your mama
'And I am your mama'
b. a ovo je s(l)onić (Antonija 2;1.17)
and this is-CLIT elephant-DIM
'And this is a small elephant'

(31) nije bocboc kaze(t)ofon (Antonija 2;0.2)
not-is yucky tape recorder
'The tape recorder is not yucky'

4.3.4. Subjects

Although Antonija is acquiring a language which allows null subjects, she uses both, overt pronominal and null subjects at the same time. The rate of lexical subjects[20] is strikingly low compared to the number of pronominal subjects, cf. Table 11. Among pronominal subjects the most frequent pronoun is *to* 'that'. The extensive use of pronouns among overt subjects is not connected to copular constructions only but also appears with other types of verbs (cf. Katičić 1997; Katičić and Schaner-Wolles 2001).

Table 11. Subject type in copular utterances, 17 one-word utterances not included, Antonija07-23 (1;7.2 - 2;1.28),(absolute numbers and row percentages)

	overt subject		
null subject	*pronominal*	*lexical*	*total*
31 (38.8%)	39 (48.8%)	10 (12.5%)	80 (100%)

5. CONCLUSION

The present study focuses on the acquisition of a specific verbal element, namely the copula, in predicative constructions in a cross-linguistic perspective (English, German, Croatian). The results not only bear on the role of the verb in early language acquisition but also on the split between languages with a root infinitive phenomenon during grammar acquisition (English, German), and languages without (Croatian).

Although adult English, German, and Croatian share the grammatical characteristic of an obligatory copula, the three languages differ in this respect in our language acquisition data. In both English and German, there is a period of linguistic development where target-like copula constructions are observed alongside copulaless predication structures that are not target-like, but compatible with universal grammar. In contrast, we find no copulaless predication structures during the acquisition of Croatian.

During the acquisition of English and German, for all the children investigated in Becker (2000a, 2000b) and in the present study, the occurrence rate of the copula is different for different kinds of predicates. Nominal predicates (temporally unbounded predicates) favor the occurrence of a copula, locative predicates (temporally bounded predicates) do not. In our investigation of two children acquiring German, we found further differences between the two clause types with and without a copula. Functional material is mostly used in the clause type with a copula. This tendency is particularly strong for pronominal subjects. Word order in clauses with a copula follows the regularities of adult verb second clauses, i.e. topicalization of the predicate into the preverbal position takes place in a substantive number of cases. Word order in predicative clauses without a copula is variable: the majority of utterances are subject initial, but subject-predicate inversion is available as well, although less frequent than the predicate initial clauses in verb second clauses containing a copula. Any structural analysis of copulaless predication will have to account for this fact.

Besides the similar distribution of predicate types with respect to the presence or absence of the copula, we find another similarity between the acquisition of the copula in English and German: the copula, if present, is

almost always finite. This is in contrast with most other verbs, which alternate between finite and non-finite forms (root infinitives) during this period of linguistic development, in both English and German. The absence of copulaless predication structures from Croatian acquisition coincides with the absence of root infinitives from Croatian acquisition. The whole picture is perfectly compatible with the hypothesis that copulaless predication structures are related to the root infinitive phenomenon. Most probably, both relate to the expression of finiteness. The finitely inflected copula as well as the finite inflection on other verbs express tense and agreement features. In case of the absence of these grammatical markers, either root infinitives or copulaless predication structures appear. It is an empirical question whether the presence or absence of root infinitives and the occurrence or non-occurrence of copulaless predication structures also cluster together in the acquisition of other languages. We leave this question open for further research.

CHRISTINE CZINGLAR
University of Vienna

ANTIGONE KATIČIĆ
University of Vienna

KATHARINA KÖHLER
University of Vienna

CHRIS SCHANER-WOLLES
University of Vienna

6. ACKNOWLEDGMENTS

We thank Hubert Haider and various audiences, in particular the participants of the conference on 'Acquisition of Verb Grammar and Verb Arguments', ZAS Berlin, 15.-17.11.2001. We are indebted to an anonymous reviewer and to Marilyn Vihman for their insightful suggestions and comments. And last but not least we would like to thank the editors, Natalia Gagarina and Insa Gülzow, for their comments and all their efforts in publishing this book.

7. NOTES

[1] The authors' names are in alphabetical order. This research was funded by the Austrian Science Fund (FWF) project P 13371-SPR 'First Language Acquisition of Austrian German in a Crosslinguistic Perspective' to Chris Schaner-Wolles. We thank the children and their parents for their patient co-operation. Thanks are also due to the Department of Speech- and Language Pathology, University of Zagreb, where the Antonija Corpus was collected as part of the project 'Psycholinguistic Aspects of the Croatian Language' (principal investigator: Marta Ljubešić). The Antonija data are also analyzed in the project 'Jezična obrada u hrvatskome: psiholingvistički i neurolingvistički pristup' (principle investigator: Melita Kovačević), funded by The Croatian Ministry of Science and Technology.

[2] Throughout this paper, we use the term 'copula omission', 'omission of the copula' as a shorthand notation for predication structures without a copula. Note that we do not make a commitment with respect to an actual omission process.

[3] Case assignment also plays an important role here, but we will not go into Case issues here.

[4] In certain stylistically marked contexts the copula can be omitted. In German the copula can be omitted in exclamatives like *Peter ein Arzt!?* 'Peter a doctor?!'. In Croatian the verb *biti* can be omitted in certain contexts, e.g. in idioms such as *Sve uzalud.* 'Everything in vain.' (cf. Katičić 1991; and Barić et al. 1997).

[5] The generative tradition uses the metaphor of trees for binary branching hierarchical structures that represent the relationships between the items in a sentence. In this terminology, the hierarchically highest node, where the tree originates, is called root. "Root clause" is the label given to a clause that is not embedded under any other clause, and translates to matrix clause or main clause in more traditional terminology.

[6] To indicate the target forms of the children's productions, the following CHILDES conventions are used, e.g. *mik [: milk]* or *mi(l)k*. Examples are glossed on a word by word basis, where spaces separate words. Typographic signs connect different parts of within-word information.

[7] Regarding modals in German, the Nico data indicate that this generalization is subject to interindividual variation: Nico uses modal verbs in the infinitive form to a substantial extent, whereas other children, like Paul, do not. Both boys are investigated here.

[8] In particular, possible distinctions inside the class of lexical verbs have been considered, cf. different opinions in Wijnen (1998); Hoekstra and Hyams (1998); Schaner-Wolles (2000); Katičić et al. (2000).

[9] Examples (8a) through (10a) are from Becker (2000a, 2000b), example (10b) is quoted from Borer and Rohrbacher (1998).

[10] A caveat regarding the term 'Aspectual Phrase' is in place here. In the following, we will not attempt to apply the exact details of Becker's analysis to our own rendition of the German and Croatian data below. Note that in particular we remain neutral with respect to the syntactic implementation of (im)perfective aspect, for which every Croatian verb is specified, and we do not tackle the question whether the AspP associated with the presence of an event argument in Becker's account is or is not a suitable locus for the perfective-imperfective distinction in Croatian.

[11] The introduction of event arguments into grammatical representations is based on considerations that not only the arguments involved in a proposition as participants play a role in the representation of the meaning of a sentence, but that events are entities that have to be represented as well.

[12] Although a general analysis of the root infinitive phenomenon is not the central subject of Becker's thesis, she considers an extension of her analysis to main verbs, cf. Becker (2000a:140ff.).

[13] Note that the account, without further assumptions, does not explain why the attested copula rate with individual-level predicates is below 100%. Apart from obvious performance factors, one factor invoked by Becker is the mapping problem between real world knowledge about permanent and temporary properties and the eventually language specific linguistic division between stage and individual predicates, cf. Becker (2000a:211f.).

[14] According to our working definition for the purpose of this paper the copula does not appear with another verbal element. We do not want to get into the discussion of the so-called 'Zustandspassiv' in German.

[15] The target language under discussion hardly allows bare nouns, even mass nouns take indefinite determiners. Proper names for persons behave exactly like descriptive noun phrases with respect to article placement. Hence they are counted as lexical noun phrases here.

[16] A similar distribution is observed in Becker's data for English, but she argues that this is due to certain discourse contexts (answers to 'what is'-questions). In our data, the amount of pronominal subjects cannot be attributed to such a discourse effect.

[17] The copula is introduced in present tense indicative only since the child did not produce other forms in the course of the analyzed period.

[18] In the variety spoken in Zagreb.

[19] The nature of clitics in Croatian constitutes a part of the general discussion on the status of clitics in general and Slavic clitics in particular with respect to phonology, morphology and syntax; cf. e.g. Ćavar and Wilder (1999); Wilder and Ćavar (1994).

[20] 'Lexical' subjects comprise bare nouns, adjectives plus nouns etc.

8. APPENDIX A

Table 12. Overview of analyzed data for Nico, predication with and without copula

Nico	*file information*			*predication cases*		
data-points	*age (y;mm.dd)*	*filename*	*duration of recording*	*cop.*	*no cop.*	*total*
1	2;3.30	Nico06	60 min.	7	41	48
2	2;4.10	Nico07	60 min.	19	15	34
3	2;4.20	Nico08	60 min.	7	30	37
4	2;5.1	Nico09	60 min.	40	16	56
5	2;5.11	Nico10	60 min.	18	10	28
6	2;5.22	Nico11	60 min.	13	21	34
7	2;6.1	Nico12	60 min.	28	17	45
8	2;6.12	Nico13	60 min.	11	22	33
9	2;6.22	Nico14	60 min.	11	32	43
10	2;7.3	Nico15	60 min.	27	21	48
11	2;7.13	Nico16	60 min.	28	19	47
12	2;7.24	Nico17	60 min.	24	54	78
13	2;8.7	Nico18	60 min.	29	14	43
14	2;8.21	Nico19	60 min.	41	9	50
15	2;9.4	Nico20	60 min.	69	7	76
total			15 hrs.	372	328	700

Table 13. Overview of analyzed data for Paul, predication with and without copula

Paul	*file information*			*predication cases*		
data-points	*age (y;mm.dd)*	*filename*	*duration of recording*	*cop.*	*no cop.*	*total*
1	2;4.6	Paul09	45 min.	3	5	8
2	2;4.20	Paul10	45 min.	1	11	12
3	2;5.16	Paul11	255 min	29	19	48
4	2;5.23	Paul12	60 min.	7	1	8
5	2;7.15	Paul13	45 min.	18	2	20
6	2;8.21	Paul17	45 min.	11	1	12
total			8 h. 15 min.	69	39	108

APPENDIX B

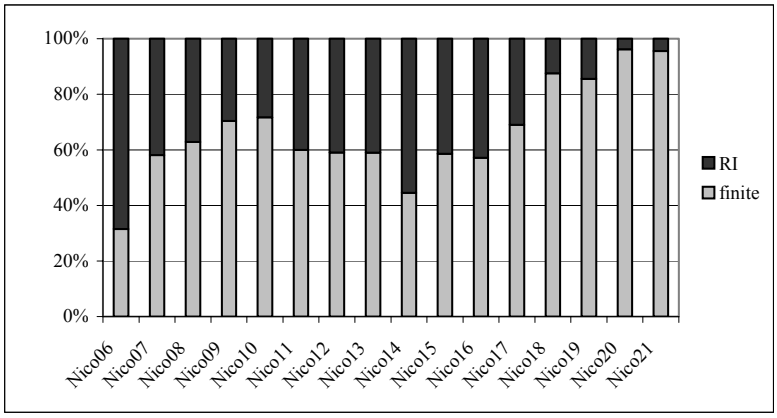


Figure 1. Ratio of finite verbs and root infinitives (RI), Nico06-21 (2;3.30 - 2;9.18), all verbs. Information on Nico06-21 is taken from Köhler and Bruyère (1995/96) and Schaner-Wolles (2000)

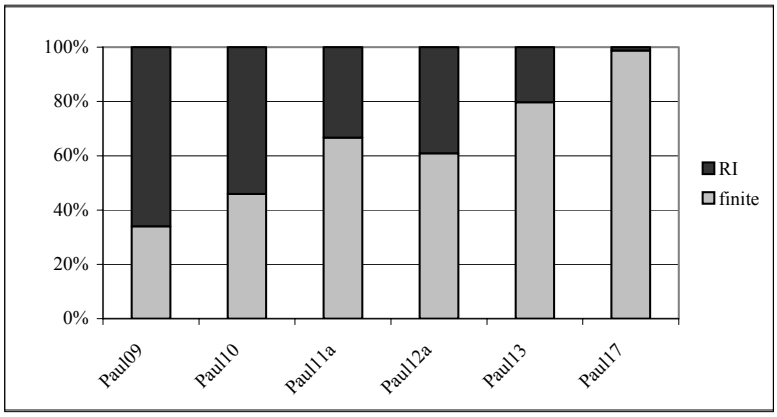


Figure 2. Ratio of finite verbs and root infinitives (RI), Paul09-17 (2;4.6 - 2;8.21), all verbs. For information on Paul09-13 we thank Katharina Korecky-Kröll (p.c.).

APPENDIX C

Table 14. Absolute numbers and relative frequencies of copulaless predication structures out of total predication structures for each type of predicate and for all types of predicates together, Nico (2;3.30 - 2;9.4)

PR	*adjectival*		*locative*		*nominal*		*other*		*all predicates*	
File	*no cop /total*	*%*	*no cop /total*	*%*	*no cop /total*	*%*	*no cop /total*	*%*	*no cop /total*	*%*
Nico06	3/3	100.0	18/19	94.7	11/17	64.7	9/9	100.0	41/48	85.4
Nico07	1/2	50.0	9/9	100.0	2/18	11.1	3/5	60.0	15/34	44.1
Nico08	3/3	100.0	17/17	100.0	2/9	22.2	8/8	100.0	30/37	81.1
Nico09	3/6	50.0	7/11	63.6	4/32	12.5	2/7	28.6	16/56	28.6
Nico10	1/1	100.0	5/9	55.6	2/14	14.3	2/4	50.0	10/28	35.7
Nico11	4/5	80.0	12/17	70.6	2/8	25.0	3/4	75.0	21/34	61.7
Nico12	3/3	100.0	8/9	88.9	5/31	16.1	1/2	50.0	17/45	37.8
Nico13	7/10	70.0	11/13	84.6	3/9	33.3	1/1	100.0	22/33	66.7
Nico14	4/4	100.0	23/25	92.0	4/13	30.8	1/1	100.0	32/43	74.4
Nico15	2/2	100.0	12/21	57.1	7/24	29.2	0/1	0.0	21/48	43.8
Nico16	0/2	0.0	15/20	75.0	3/24	12.5	1/1	50.0	19/47	40.4
Nico17	7/9	77.8	21/28	75.0	20/35	57.1	6/6	100.0	54/78	69.2
Nico18	4/5	80.0	7/12	58.3	3/24	12.5	0/2	0.0	14/43	32.6
Nico19	1/8	12.5	1/7	14.3	7/31	22.6	0/4	0.0	9/50	18.0
Nico20	3/36	8.3	2/12	16.7	2/25	8.0	0/3	0.0	7/76	9.2
total	46/98	46.9	168/229	73.4	77/314	24.5	37/59	62.7	328/700	46.9

Table 15. Absolute numbers and relative frequencies of copulaless predication structures out of total predication structures for each type of predicate and for all types of predicates together, Paul (2;4.6 - 2;8.21)

PR	*adjectival*		*locative*		*nominal*		*other*		*all predicates*	
File	*no cop /total*	*%*	*no cop /total*	*%*	*no cop /total*	*%*	*no cop /total*	*%*	*no cop /total*	*%*
Paul09	1/2	50.0	2/2	100.0	2/2	100.0	0/2	0.0	5/8	62.5
Paul10	2/2	100.0	7/7	100.0	1/1	100.0	1/2	50.0	11/12	91.7
Paul11	3/5	60.0	9/19	47.4	1/16	6.25	6/8	75.0	19/48	39.6
Paul12	0/1	0.0	1/5	20.0	0/0		0/2	0.0	1/8	12.5
Paul13	2/7	28.6	0/8	0.0	0/5	0.0	0/0		2/20	10.0
Paul17	0/4	0.0	1/5	20.0	0/3	0.0	0/0		1/12	8.33
total	8/21	38.1	20/46	43.5	4/27	14.8	7/14	50.0	39/108	36.1

APPENDIX D

Table 16. Overview of analyzed data for Antonija07-23 (1;7.2 - 2;1.28)

file information				*predication cases*		
data-points	*age (y;mm.dd)*	*filename*	*number of utterances*	*cop.*	*no cop.*	*total*
1	1;7.2	antbla07	79	10	0	10
2	1;7.15	antbla08	141	9	0	9
3	1;7.27	antbla09	101	2	0	2
4	1;9.15	antbla10	207	24	0	24
5	1;10.0	antbla11	152	10	0	10
6	1;10.10	antbla12	57	7	0	7
7	1;10.21	antbla13	74	2	0	2
8	1;10.30	antbla14	55	1	0	1
9	1;11.10	antbla15	34	0	0	0
10	1;11.17	antbla16	60	2	0	2
11	1;11.25	antbla17	77	1	0	1
12	2;0.2	antbla18	90	2	0	2
13	2;0.12	antbla19	93	2	0	2
14	2;0.28	antbla20	53	1	0	1
15	2;1.8	antbla21	35	3	0	3
16	2;1.17	antbla22	238	19	0	19
17	2;1.28	antbla23	97	2	0	2
total			1643	97	0	97

9. REFERENCES

Anđel, Maja, Sabine Klampfer, Marianne Kilani-Schoch, Wolfgang U. Dressler, and Melita Kovačević. "Acquisition of verbs in Croatian, French and Austrian German – an outline of a comparative analysis." *Suvremena lingvistika* 49-50.1-2 (2000): 5-25.

Barić, Eugenija, Mijo Lončarić, Dragica Malić, Slavko Pavešić, Mirko Peti, Vesna Zečević, and Marija Znika. *Hrvatska gramatika*. Zagreb: Školska knjiga, 1997.

Becker, Misha. The Development of the Copula in Child English: The Lightness of Be. PhD Dissertation. University of California, Los Angeles, 2000a.

Becker, Misha. "The Acquisition of the Copula in Child English." In Roger Billerey and Danielle Lillehaugen (eds.), *WCCFL 19 Proceedings*. Somerville, MA: Cascadilla Press, 2000b: 57-70.

Bloom, Lois. *Language Development: Form and function in emerging grammars*. Cambridge, MA: MIT Press. 1970.

Borer, Hagit and Bernhard Rohrbacher. "Minding the Absent: Arguments for the Full Competence Hypothesis." Ms., University of Southern California, Los Angeles and Northwestern University, Chicago, 1998.

Brown, Roger. *A first language*. Cambridge, MA: Harvard University Press, 1973.

Cardinaletti, Anna and Maria Teresa Guasti (eds.). *Small Clauses*. *Syntax and Semantics 28*. London: Academic Press, 1995.

Carlson, Gregory. Reference to Kinds in English. PhD Dissertation. University of Massachusetts, Amherst, 1977.

Carlson, Gregory. *Reference to Kinds in English*. New York: Garland Publishing, 1980.

Ćavar, Damir and Chris Wilder. "'Clitic third' in Croatian." In Henk van Riemsdijk (ed.), *Clitics in the languages of Europe*. Berlin: de Gruyter, 1999: 429-467.

Davidson, Donald. "The Logical Form of Action Sentences." In Nicholas Rescher (ed.), *The Logic of Decision and Action*. Pittsburgh: The University Press, 1967: 81-95.

Déchaine, Rose-Marie. Predicates across categories. Towards a category-neutral syntax. PhD Dissertation. University of Massachusetts, Amherst, 1993.

Enç, Mürvet. "Anchoring Conditions for Tense." *Linguistic Inquiry* 18.4 (1987): 633-657.

Guasti, Maria Teresa. *Language Acquisition: The Growth of Grammar*. Cambridge, MA: MIT Press, 2002.

Hoekstra, Teun and Nina Hyams. "Aspects of root infinitives." *Lingua* 106 (1998): 81-112.

Katičić, Antigona and Chris Schaner-Wolles. "Subjekt i predikat u usvajanju hrvatskoga jezika." Paper presented at the *15th Annual Conference of the Croatian Applied Linguistics Society (Hrvatsko društvo za primjenjenu lingvistiku) on Psycholinguistics and Cognitive Science in Croatian Applied Linguistics.*, Opatija, May 2001.

Katičić, Antigone. "Zum Erstspracherwerb des Kroatischen: Morphologische und syntaktische Aspekte beim Erwerb des Verbalsystems." M.A. Thesis. University of Vienna, 1997.

Katičić, Antigone. "Early Verb Development in one Croatian-speaking Child." In Dagmar Bittner, Wolfgang U. Dressler and Marianne Kilani-Schoch (eds.), *First Verbs: On the way to mini-paradigms.* ZAS Papers in Linguistics 18. Berlin: ZAS, 2000: 115-129.

Katičić, Antigone, Sabine Klampfer, Katharina Korecky-Kröll, and Katharina Köhler. "Grammatische Verben im Erstspracherwerb: Österreichisches Deutsch im Vergleich zu Kroatisch." Paper presented at the *28th Austrian Linguistics Conference*, University of Graz, December 2000.

Katičić, Radoslav. *Sintaksa hrvatskog književnog jezika*. Zagreb: Globus and Hazu, 1991.

Köhler, Katharina and Sabine Bruyère. "Finiteness and verb placement in the L1-acquisition of German." *Wiener Linguistische Gazette* 53-54 (1995/96): 63-83.

Köhler, Katharina. "Finitheit und Verbposition. Eine Fallstudie zum Erstspracherwerb des Deutschen." MA Thesis. University of Vienna, 1998.

Korecky-Kröll, Katharina. "Der Erwerb von Subjekt- und Objektklitika in der Sprachentwicklung eines Wiener Kindes im Alter von 2-3½ Jahren." MA Thesis. University of Vienna, 2000.

Koster, Jan. "Dutch as an SOV-language." *Linguistic Analysis* 1 (1975): 111-136.

Kratzer, Angelika."Stage-Level and Individual-Level Predicates." Ms. University of Massachusetts, Amherst, 1988.

Kratzer, Angelika. "Stage-Level and Individual-Level Predicates." In Gregory Carlson and Francis J. Pelletier (eds.), *The Generic Book*. Chicago: University of Chicago Press, 1995: 125-175.

Löbel, Elisabeth. "Copular Verbs and Argument Structure." Paper presented at the *Workshop on Predication,* University of Trondheim, October 2001.

MacWhinney, Brian and Catherine Snow. "The child language data exchange system." *Journal of Child Language* 17 (1990): 457-472.

Magner, Thomas F. *A Zagreb Kajkavian Dialect*. The Pennsylvania State University Studies 18. University Park: Pennsylvania State University Press, 1966.

Poeppel, David and Ken Wexler. "The full competence hypothesis of clause structure in early German." *Language* 69.1 (1993): 1-33.

Radford, Andrew. *Syntactic theory and the acquisition of English syntax: The nature of early child grammar of English*. London: Blackwell, 1990.

Rothstein, Susan D. "Three Forms of English *Be*." *MIT Working Papers in Linguistics* 9 (1987): 225-238.

Schaner-Wolles, Chris. "Intermodular Synchronization: On the Role of Morphology in the Normal and Impaired Acquisition of a Verb-Second Language." In Rosemarie Tracy and Elsa Lattey (eds.), *How tolerant is universal grammar? Essays on language learnability and language variation*. Tübingen: Niemeyer, 1994: 205-224.

Schaner-Wolles, Chris. "The Acquisition of Negation in a Verb-Second Language." *Wiener Linguistische Gazette* 53-54 (1995/96): 87-119.

Schaner-Wolles, Chris. "Am Anfang ist das Verb, finit markiert ist es erst später." Paper presented at the *Annual Conference of the German Linguistics Society (Deutsche Gesellschaft für Sprachwissenschaft),* University of Marburg, March 2000.

Stowell, Tim. "What was there before there was there?" In Donka Farkas, Wesley M. Jacobson, and Karol W. Todrys (eds.), *Papers from the Fourteenth Regional Meeting of the Chicago Linguistic Society*. Chicago: University of Chicago, 1978: 458-471.

Stowell, Tim. "Subjects across categories." *The Linguistic Review* 2 (1983): 285-312.

Wexler, Ken. "Optional infinitives, head movement, and the economy of derivations." In David Lightfoot and Norbert Hornstein (eds.), *Verb Movement*. Cambridge: Cambridge University Press, 1994: 305-350.

Wijnen, Frank. "The temporal interpretation of Dutch children's root infinitivals: The effect of eventivity." *First Language* 18 (1998): 379-402.

Wilder, Chris and Damir Ćavar. "Long Head Movement? Verb Movement and Clitization in Croatian." *Lingua* 93 (1994): 1-58.

Williams, Edwin. "Predication." *Linguistic Inquiry* 11 (1980): 203-238.

Williams, Edwin. "There-Insertion." *Linguistic Inquiry* 15 (1984): 131-153.

Williams, Edwin. "The Asymmetry of Predication." In Ralph Blight and Michelle Moosally (eds.), *Proceedings of the 1997 Texas Linguistic Society Conference*. Texas Linguistic Forum 38. Department of Linguistics, University of Texas at Austin 1997: 323-333.

CHILDES – Child Language Data Exchange System: http://www.childes.psy.cmu.edu

NATSUKO TSUJIMURA

WHY NOT ALL VERBS ARE LEARNED EQUALLY

The Intransitive Verb Bias in Japanese[*]

Abstract. This paper investigates whether children learn all verbs with equal ease. Based on a longitudinal study of a Japanese-speaking child, I will demonstrate that the child has a bias toward intransitive verbs. Moreover, intransitive verbs that enter into the child's early vocabulary tend to be verbs that denote motion with specific path or definable goal. It is claimed that a driving force for the lexical bias is a language-specific principle that focuses on result, as opposed to process, of an action. The conclusion lends support to the view that language-specific principles can guide the child in the development of verb semantics and conceptualization of events.

1. INTRODUCTION

Much recent research in child language acquisition focuses on the acquisition of different parts of speech. In this pursuit, examinations of verb acquisition have raised a number of intriguing issues in the fields of linguistics and psychology. One of these issues includes the question of lexical bias. Concerned with the question of the acquisition order among various lexical classes and its conceptual-perceptual motivation, Gentner (1982) in her proposal of the natural partitions hypothesis discusses children's noun bias in their lexical acquisition. According to the natural partitions hypothesis, the proportion of nouns in children's vocabulary is greater than that of predicate terms because nouns are conceptually and perceptually simpler and less relational than predicative words like verbs and prepositions (cf. Herr-Israel and McCune, this volume). As Gentner and Boroditsky (2001) most recently explain, a bias toward a particular word category or within a single word category can be motivated by cognitive and linguistic dominance.[1] While it seems to have been assumed by many that developmentally cognitive dominance precedes linguistic dominance, the

N. Gagarina and I. Gülzow (eds.), The Acquisition of Verbs and their Grammar: The Effect of Particular Languages, 105–122.

opposite direction and bi-directional relationship have also been emphasized (Smith and Sera 1992; Choi 1997).

In this paper I will consider the question of why intransitive verbs of a certain kind seem to be easier for Japanese children to learn, and discuss a language-specific principle that may lead to an answer to this question. More specifically, I wish to demonstrate that the input language leads the child to a principle that focuses on the result, rather than the process, of an event, and that such a principle creates a feeding environment for children's development of verb semantics and conceptualization of events (Choi and Bowerman 1991; Bowerman 1996a, b; Choi 1997). My discussion will be based on the examination of longitudinal production data of a Japanese monolingual child, between the ages of 1;6 to 3;0.

2. SOME CHARACTERISTICS OF JAPANESE

I will begin by describing some characteristics of Japanese that are relevant in this paper. The Japanese input language minimally provides syntactic information as to whether an overt NP is an argument or an adjunct of a verb, how many argument NPs a verb is subcategorized for, and which grammatical function an argument NP is supposed to bear. The lack of information of this sort is attributed to zero pronouns, Case drop, scrambling, and any combination of these, as illustrated in the examples given in (1-4).

(1) zero pronouns
   a. Taroo-ga tegami-o kaita.
      Taro-Nom letter-Acc wrote
      'Taro wrote a letter.'
   b. Taroo-ga kaita.
   c. Tegami-o kaita.
   d. Kaita.

(2) Case drop
   Taroo(-ga) tegami(-o) kaita.
   Taro(-Nom) letter(-Acc) wrote
   'Taro wrote a letter.'

(3) scrambling
Tegami-o Taroo-ga kaita.
letter-Acc Taro-Nom wrote
'Taro wrote a letter.'

(4) combination
a. Tegami(-o) Taroo(-ga) kaita.
letter(-Acc) Taro(-Nom) wrote
b. Tegami(-o) kaita.
letter(-Acc) wrote

Examples in (1-4) are normal input sentences which any adult speaker would perceive as grammatical. (1a) instantiates the canonical order with all the verb's subcategorized arguments expressed overtly. The subject is normally marked with the Nominative Case marker *–ga* and the object with the Accusative Case marker *–o*. In (1b-d) the object, the subject, and both arguments are missing, but these zero pronouns are properly interpreted contextually. The Nominative and Accusative Case markers indicate the grammatical function of the NPs to which they are suffixed, but they are optional, as is illustrated in (2), without changing the interpretation that is conveyed in (1a). The scrambling phenomenon, another optional operation, allows permutation of arguments and adjuncts with a large amount of possible variation. This is shown in (3). The phenomena of zero pronouns, Case drop, and scrambling can appear simultaneously, as (4) illustrate.[2]

In English, none of the options of (1-4) is possible: the application of these phenomena based on the canonical sentence of (5a) leads to ungrammatical sentences of (5b-g).

(5) a. Mary made a cake.
b. *Mary made.
c. *Made a cake.
d. *A cake made Mary.
e. *Made Mary.
f. *A cake made.
g. *Made.

3. TRANSITIVITY PARADIGM

While the transitivity distinction can hardly be drawn from the syntactic information given the phenomena described above, it is reflected in the morphology of a particular set of verb pairs, namely, morphologically related intransitive-transitive verb pairs. In English an intransitive verb and its transitive counterpart that establish the inchoative vs. causative dichotomy mostly take an identical morphological form, with a few exceptions including verb pairs like *rise-raise* and *lie-lay*. Examples of this sort are observed in *break, freeze, melt, bend, open, close*, and *sink*, among many more, as in "They broke the window," vs. "The window broke," etc. In contrast, Japanese employs a number of intransitive-transitive verb pairs that share identical roots, and to these roots different suffixes can be attached for the purpose of transitivity distinction However, these intransitive- and transitive-forming suffixes are not unique. Furthermore, in some cases, an intransitive-forming suffix takes the same morphological make-up as a transitive-forming suffix, and there is very little regularity in determining which verb takes which morpheme. Jacobsen (1992) lists over a dozen such morpheme pairs to describe these morphologically related intransitive-transitive verb paradigm as in (6). Some examples are given in (7) (The last morpheme in (7), *–(r)u,* marks the present tense.)

(6) intransitive/transitive-forming morphemes (Jacobsen 1992)

-e-/-@-	-e-/-as-	(-@- = zero morpheme)
-i-/-as-	-@-/-se-	
-@-/-e-	-@-/-as-	
-ar-/-e-	-ri-/-s-	
-ar-/-@-	-e-/-akas-	
-r-/-s-	-or-/-e-	
-re-/-s-	-are-/-e-	
-i-/-os-	and miscellaneous pairs	

(7)

intransitive	transitive	gloss
ak-u	ak-e-ru	'open'
kowa-re-ru	kowa-s-u	'break'
nig-e-ru	nig-as-u	'escape'
nob-i-ru	nob-as-u	'extend'
ag-ar-u	ag-e-ru	'rise'
ok-i-ru	ok-os-u	'get up'

The semantic classes of verb pairs that enter into this paradigm include, but are not limited to, change of location/position and change of state. As I will discuss in the next section, the semantic notion of change of location/position plays an important role in the child's acquisition of intransitive verbs.

4. RESULTS AND DISCUSSION

I examined a longitudinal corpus of a Japanese monolingual child reported in Miyata (1993) and Miyata (1995) (Miyata 1992; MacWhinney 1995; Oshima-Takane and MacWhinney 1998). The child is a boy, Aki, and the observation period was from 1;6 to 3;0. The numbers of morphologically related transitive-intransitive verb pairs that appeared in Aki's speech and in the input language were counted. Also counted were morphologically unmarked intransitive verbs, for the reasons that I explain in section 4.3. I have excluded frequent instances of *suru* 'do' and *yaru* 'do' because they serve as pro-verbs replacing verbs of any dynamic event type, thereby often making the transitivity distinction unclear. I have also excluded repeated words from the token count. In what will follow below, I shall illustrate a bias toward intransitive verbs and will discuss a possible explanation for it.

4.1. Frequent Use of Morphologically Marked Intransitive Verbs

Let us first compare morphologically marked intransitive verbs and their transitive counterparts in Aki's production data, focusing on the token count. The observation period is divided into the following five developmental stages.

I: 1;6-1;11
II: 2;0-2;2
III: 2;3-2;5
IV: 2;6-2;8
V: 2;9-3:0

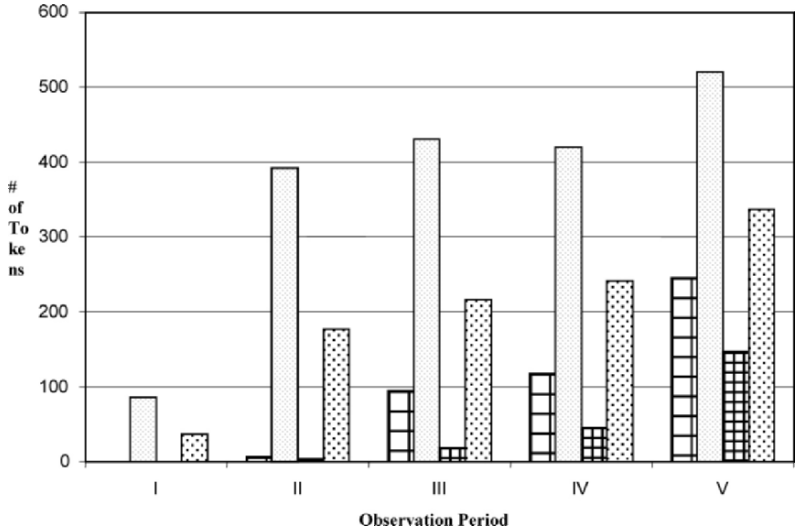


Figure 1. Morphologically Marked Iintransitive (Vi) and Transitive (Vt) Verbs – Aki and Input [Token]. For each observation period, the four bars indicate from left to right (i) Aki's Vi, (ii) Input's Vi, (iii) Aki's Vt, and (iv) Input's Vt.

Figure 1 illustrates that in morphologically related intransitive-transitive verb pairs, Aki produced more intransitive verb tokens than their transitive counterparts throughout the observation period except the first. Furthermore, the dominance in morphologically marked intransitive verb production is the pattern that is observed in the input language as well. The intransitive verb dominance observed in morphologically related transitive-intransitive verb pairs is consistent with Nomura and Shirai (1997). The same tendency in input language is also observed by Nomura and Shirai although their observation period is shorter (1;5-2;1) than the one analyzed here. Based on their observation, Nomura and Shirai (1997) conclude that the intransitive dominance in the child's production reflects the input distribution pattern of these verb pairs. It still remains to be answered, however, as to what might be the exact nature of the child's intransitive verb preference. It is certainly not the case that all the lexical patterns of input are mirrored in child language. For instance, a comparison between the number of

tokens and types of nouns and verbs that Aki's speech and that of his input demonstrates that while the input contained more verb tokens but more noun types, Aki produced more nouns in both token and type. As Gentner and Boroditsky (2001) argue, input may *influence* children's lexical acquisition patterns, but it does not *account for* specific behavior they exhibit.

In exploring the exact nature of the intransitive bias, Slobin's (1991) experiment on narrative development is relevant. In his cross-linguistic study of narrative development, which includes 9 different languages, Slobin (1991) found language-specific linguistic patterns of describing scenes that are pertinent to rhetorical style. According to Slobin, for example, what is asserted and what is implied in a certain structure is different depending on the specific language: English asserts trajectory while implying resultant locative states, and Spanish asserts locations and directions while implying trajectory. Slobin gives the examples in (8-9) to illustrate this difference (Slobin 1991:19).

(8) English: Assert trajectory, imply end-state.
The boy climbed the tree.

(9) Spanish: Assert end-state, imply trajectory.
El nino esta subido en el arbol.
'The boy is climb-PART en [=in/on] the tree.'
[=the boy is in a state of having climbed the tree]

Slobin claims that this differentce has an implication to overall rhetorical style, and that in their narrative English speakers focus more on the process while Spanish speakers on the result. Slobin's point is that when the child acquires the grammar, s/he learns these language-specific patterns encoded in the language that are relevant to discourse.

I wish to claim that the Japanese transitivity paradigm encodes a principle on perspective differences, and that Aki's higher number of morphologically marked intransitive verb tokens than their transitive counterparts reflects this principle on perspectives. The majority of morphologically related intransitive and transitive verb pairs in Japanese denote change of state or change of location/position. Furthermore, transitive verbs of these pairs are lexical causatives and are contrasted with productive causatives that are formed by suffixation of the morpheme *–(s)ase* to a verb stem. The major difference between these two kinds of causatives which has been noted is that lexical causatives denote direct

causation while productive causatives imply indirect causation (Shibatani 1990). For example, the lexical causative verb, *tome-ru* 'stop', is contrastive with its productive causative counterpart, *tomara-se-ru* 'make X stop': the latter is derived by suffixing the causative morpheme *–(s)ase* to the intransitive verb *tomar(u)*. According to Shibatani's (1990: 317) explanation, the lexical causative *tomeru* "expresses manipulative causation in which the causer brings about the caused event by physically manipulating the cause," while the productive causative counterpart *tomaraseru* "expresses directive causation in which the causer gives a direction to the cause to bring about the caused event." Thus, the causative event scene that is captured by morphologically related intransitive-transitive verb pairs is consistent with what Slobin (1985) calls "manipulative activity scene," which is a prototypical event encoded in transitive clauses. Direct causation events denoted by the verb pairs in question, as a prototypical scene, naturally receive Aki's attention. Given the input language that exhibits heavier use of intransitive verbs, the child focuses on results as a way to describe the prototype scene of direct causal events. This is why Aki produced more morphologically marked intransitive tokens than their transitive counterparts, assuming that the token count reflects the frequency of verb use after he has learned it. Furthermore, Aki maintains the heavier use of intransitive verbs as rhetorical style of his language: this is reflected by the fact that the intransitive bias is maintained throughout the observation period. The rhetorical phenomenon of emphasizing result in adult language has long been noted by many scholars of Japanese (Ikegami 1981; Teramura 1976; Kageyama 1996, etc.) as is distinctively demonstrated in translation. Ikegami (1981), for example, attributes the rhetorical phenomenon to typological differences between "DO-languages" and "BECOME-languages": English belongs to the former, and Japanese to the latter. Roughly speaking, DO-languages tend to describe events focusing on an agent's action while BECOME-languages tend to do so focusing on a resulting situation. It should be pointed out that there are a few other morphological mechanisms to encode the result-focused perspective in the language.[3]

4.2. Relative Ease in Learning Morphologically Marked Intransitive Verbs

We have just observed Aki's intransitive verb bias focusing on the frequency of his use of morphologically marked intransitive verbs. We have done so by examining the token count. In order to investigate the relative ease or difficulty of initial learning of these verbs, however, we now need to look at the type count. Figure 2 shows the type count of Aki's morphologically related intransitive-transitive verb pairs.

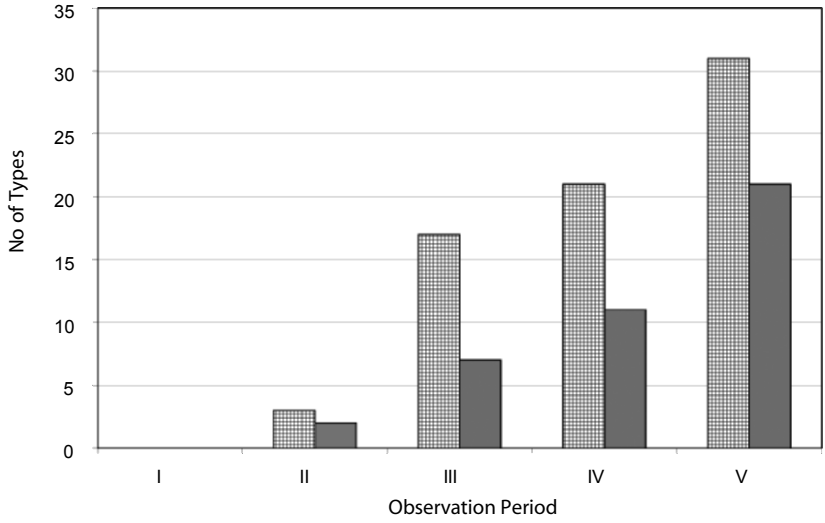


Figure 2. Morphologically Marked Intransitive (Vi) and Transitive (Vt) Verbs – Aki [Type 1]. For each observation period, the left bar indicates Vi and the right bar Vt.

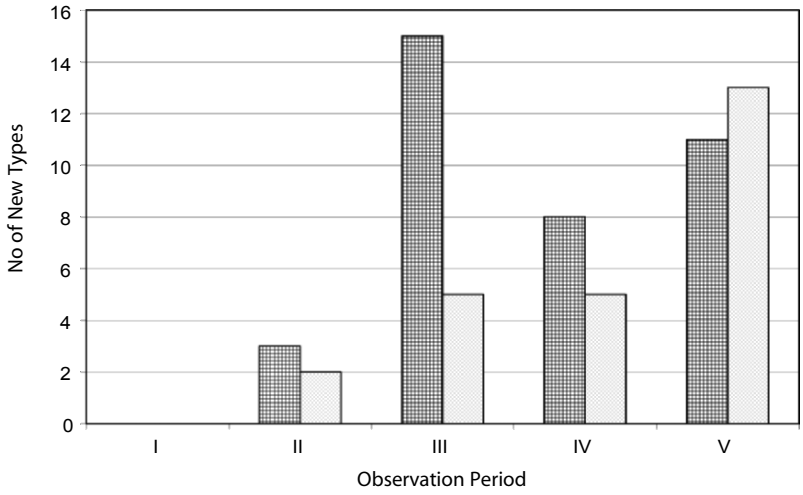


Figure 3. Morphologically Marked Intransitive (Vi) and Transitive (Vt) Verbs – Aki [Type 2]. For each observation period, the left bar indicates Vi and the right bar Vt.

"Type 2" in Figure 3 indicates the number of types that are new during that particular developmental stage, and shows that Aki increased the number of new types of morphologically marked intransitive verbs during the third developmental stage, while such a sharp increase in morphologically marked transitive verbs did not take place until toward the end of the last stage. These type counts suggest that Aki exhibits relative ease in learning morphologically marked intransitive verbs.

It should further be noted that a morphologically marked intransitive verb emerged earlier than its transitive counterpart. 43 pairs of morphologically related intransitive and transitive verb pairs have been found in Aki's corpus, and the earlier emergence of intransitive verbs was observed in 29 of the pairs.[4]

The questions that I would like to concentrate are: what is it that makes Aki's acquisition of morphologically marked intransitive verbs easier than transitive verbs? Are there any semantic generalizations to be drawn from a group of morphologically marked intransitive verbs that Aki has learned initially? I believe that the previously mentioned language-specific principle of the result-focused

perspective again plays an important role in earlier learning of morphologically marked intransitive verbs.

4.3 Interaction of [Motion] and the Result-Focused Perspective

The three morphologically marked intransitive verbs that come into Aki's vocabulary during the second developmental stage are *hairu* 'enter/be put in', *deru* 'exit/be taken out', and *oriru* 'get off/be taken down', as in (10).

(10) *hairu* 'enter/be put in' (transitive: *ireru* 'put x in')
deru 'exit/be taken out' (transitive: *dasu* 'take x out')
oriru 'get off/be taken down' (transitive: *orosu* 'take x down')

These verbs denote change of location. Their respective morphologically marked transitive verbs refer to causal events, and these three intransitive verbs describe a state as a result of the causal event denoted by their transitive counterparts. At the same time, these intransitive verbs also denote inherently directed motion. That is, these verbs are systematically polysemous in the adult grammar, as I have demonstrated in Tsujimura (2002). The polysemous situation is described in (11).

(11) a. Ano ookii piano-ga uti-ni haitta.
that big piano-Nom house-in be put in
'That big piano has been put in the house.'
[result of change in location]
(cf. John-ga piano-o uti-ni ireta.)

b. Hanako-ga uti-ni haitta.
Hanako-Nom house-in entered
'Hanako entered the house.' [inherently directed motion]

There are 15 new intransitive verb types that entered Aki's vocabulary during the third developmental stage (III in Figure 3). Of these 15 verb types, 11 denote change of location/position. They are given in (12), and crucially, all of them except the last three exhibit the same range of polysemous situation as the initial three verbs, *hairu* 'enter/be put in', *deru* 'exit/be taken out', and *oriru* 'get off/be taken down' in (10).

(12) *narabu* 'line up', *noru* 'get on', *tomaru* 'stop', *tooru* 'pass', *magaru* 'turn', *okiru* 'get up', *kuttsuku* 'attach', *ochiru* 'fall', *toreru* 'detach', *aku* 'open', *tsuku* 'attach'

Furthermore, the ratio of motion verbs to the total number of morphologically unmarked intransitive verbs in the second through the fourth developmental stages indicates that about two thirds of those verbs denote motion. The ratio is shown in (13).

(13) morphologically unmarked intransitive verbs (type)

	total	motion
II	6	5 (83%)
III	15	10 (66%)
IV	25	17 (68%)
V	35	18 (51%)

And, strikingly these morphologically unmarked intransitive motion verbs denote inherently directed motion: verbs include *iku* 'go', *kuru* 'come', *tsuku* 'arrive', *kaeru* 'return', *noboru* 'climb up', *dekakeru* 'leave', and *tobidasu* 'come out suddenly', as in (14a), and motion whose path or ending location/position (i.e., the location/position at which motion ends) can be recognizable, such as *korobu* 'trip over, fall', *suwaru* 'sit down', *hamidasu* 'come out', and *todoku* 'reach', as in (14b).

(14) examples of morphologically unmarked intransitive verbs
   a. *iku* 'go', *kuru* 'come', *tsuku* 'arrive', *kaeru* 'return', *noboru* 'climb up', *dekakeru* 'leave', *tobidasu* 'come out suddenly'
   b. *korobu* 'trip over, fall', *suwaru* 'sit down', *hamidasu* 'come out (of a boundary)', *todoku* 'reach'

The conceptual properties that tie together the verbs in (10), (12), and (14) as a natural class are motion and path or ending position. For example, inherently directed motion verbs such as *hairu* 'enter' in (10), *ochiru* 'fall' in (12), and *iku* 'go' in (14a) conflate motion and path in the sense of Talmy (1985); and verbs like *tsuku* 'attach' in (12), *toreru* 'detach' also in (12) and *hamidasu* 'come out

(of a boundary)' in (14b) do not generally denote spontaneous motion by a volitional agent, but events named by these verbs undergo motion whose ending location is specific in relation to a reference point. So, we can generalize that Aki attends more to a motion event with a terminal or reference point that can define the result of the event.

I should note that in the adult grammar, the two senses of verbs like *hairu* 'enter/be put in' and *deru* 'exit/be taken out' described in (11) exhibit different linguistic properties including the aspectual one, as I have discussed in Tsujimura (2002). However, Aki does not seem to be susceptible to such differences, and treats them in parallel. Consider his actual speech in (15). Both senses emerged around the same time.

(15) a. *deru* 'exit'/'be taken out'

2;2	(takes a train, and hold up the train)	
	Deta. 'It's out.'	[change of location]
2;3	Otoosan deta. 'Dad came out.'	[directed motion]
2;5	Densha dechau. 'The train is about to leave.'	[directed motion]
	(cf.2;1 Deta. 'I poo-pooed')	

b. *hairu* 'enter'/'be put in'

2;2	(takes a puzzle piece, and fits it in)	
	Haitta. 'It's put in.'	[change of location]
2;3	Aki: Seesoosha. 'a cleaning car'	
	Mo: Gomi doko? 'Where is garbage?'	
	Aki: Gomi koko. 'Garbage is here.'	
	Haitta. 'It's in.'	[change of location]
2;3	(talking about entering somewhere narrow)	
	Aki-chan hairu. 'I'm going to enter.'	[directed motion]

c. *noru* 'get on'/'be put on'

2;3	Kore notte! 'Get on.'	[directed motion]
2;4	(talking about putting cargo onto forklift)	
	A notta. 'It's on.'	[change of location]

What glues these two senses together is the presence of a definable terminal location, which is consistent with the generalization that has been drawn above. And, in most of the particular cases of (15), it is likely that the scenes described as change of location may be accompanied by some motion as well.

Recall that we have motivated a language specific principle on perspectives to explain the frequent use of Aki's morphologically marked intransitive verbs. This principle "trains" Aki to focus on result of an event. The result-focused perspective, then, affects Aki's conceptualization of an event. Given an event that involves motion, which is concrete and constitutes a salient concept, Aki applies the principle of perspectives, and isolates ending location/position. Morphologically marked intransitive verbs that denote change of location/position represent two salient concepts, motion and ending location/position: motion is perceptually concrete; and ending location/position provides a natural match for the result-focused perspective. Morphologically marked transitive verbs, in contrast, encode process, and since the target of focus is not consistent with the principle, they take more time to be acquired. Going back to our question, what makes Aki's acquisition of morphologically marked intransitive verbs easier than transitive verbs, the principle of result-focused perspective refers precisely to the concept that is represented in morphologically marked intransitive verbs, namely, ending location/position.[5]

5. CONCLUSION

To sum up, I have discussed Aki's acquisition bias toward morphologically marked intransitive verbs, and have argued that a language-specific principle that result, rather than process, should be focused in describing an event triggers Aki's frequent production of those intransitive verbs. This choice in perspective is encoded in the transitivity distinction in morphologically related verb pairs in Japanese, and hence the intransitive bias examined in this paper is a natural consequence of such a language-specific principle. I then claimed that the same principle plays a central role in Aki's acquisition bias toward morphologically marked intransitive verbs. Based on a relatively concrete and thus salient concept of motion, Aki applies the result-focus perspective to a motion event, which helps Aki isolate ending location/position, or more generally goal, as a coherent concept. The principle of result-focused perspective, learned from the input, first guides Aki to recognize that ending location/position is a coherent concept represented in some morphologically marked intransitive verbs, making his initial learning of these verbs relatively straightforward. Once he acquires them, he produces them more frequently than their transitive counterparts, following the same principle.

The examination of Aki's morphologically marked intransitive verb bias has several implications. First, our discussion of Aki's heavier use of morphologically marked intransitive verbs suggests that the child learns a language-specific principle from early on and maintains it. We are free to adopt any point of view in our perception of a scene, process-focused or result-focused for instance, but a particular perspective for describing an event must be learned through the language more generally rather than through perceptual experience. Aki's input, which is filled with actual instantiation of the principle, provides an excellent background for such learning. Once the child learns this principle, he chooses intransitive *open* over transitive *open*, for example, not based on its intransitive form per se but for a rhetorical reason that is consistent with the principle he learned. Furthermore, as Slobin (1991) comments, once the child is trained to adopt a particular perspective in describing a scene, it is very difficult to "retrain" him. We have seen this is the case with Aki: that is, there was not a single developmental stage at which a transitive verb bias was observed.

Second, I have suggested that the acquisition of the language-specific principle on perspectives may create a feeding environment for finding conceptual naturalness in initial verb learning. As several other papers in this volume demonstrate (e.g., Czinglar et al.; Guelzow and Gagarina, and Vihman and Vija), we can hardly downplay the significant role that individual variation among languages plays in children's verb acquisition. If the view considered here is supported in future research, it will provide evidence that linguistic, and language-specific, principles can promote a specific type of conceptual development (Bowerman 1996a, b; Choi 1997, Smith and Sera 1992).

Third, I have argued that the morphologically marked as well as unmarked intransitive verbs that Aki acquired during the earlier stages of verb learning are verbs that denote motion with specific path or definable goal. The observation we made of motion events may in fact shed light on what Tomasello (1992, 1995) calls the "packaging problem": that is, there are many ways of defining a verb given an event to which the verb refers, but the child must figure out which aspects of the event are packaged in the verb. Our discussion of motion verbs, then, indicates that a discourse-oriented principle can help the child discern the two aspects of motion events, i.e., motion and path, as conflatable components of Japanese motion verbs.

6. NOTES

I would like to thank Ken de Jong, Caitlin Dillon, Stuart Davis, Dan Dinnsen, Julien Musolino, Linda Smith, and the audience of the Ling Lunch at Indiana University for their comments and suggestions on various stages of this work. I am particularly indebted to Melissa Bowerman for her generous guidance on a number of issues that have come to be the foundation of the work reported in this paper. This work has been supported by Office of Overseas Programs and East Asian Studies Center at Indiana University through their Overseas Conference Fund and Faculty Travel Grant.

[1] Gentner and Boroditsky (2001: 215) explain the two types of dominance as follows. Cognitive dominance is the situation in which "concepts arise from the cognitive-perceptual sphere and are simply named by language." Linguistic dominance, on the other hand, is the situation where "the world presents perceptual bits whose clumping is not pre-ordained, and language has a say in how the bits get conflated into concepts."

[2] It is well documented that input language is filled with these phenomena. Nakayama (1996), for example, surveyed 1259 sentences in 4 Japanese adults' spontaneous speech, and found that 71% of them contained zero pronouns. Rispoli (1991) also reports that of 226 transitive sentences spoken by 9 Japanese caregivers, only 1% were complete sentences with subject NPs marked with Nominative Case and object NPs marked with Accusative Case. Rispoli furthermore states that in the caregivers' transitive sentences, 32% appear only with a verb and 46% occur with a verb and a single NP without Nominative or Accusative Case marker. As Nakayama (1996) reports, on the other hand, children demonstrate a high level of understanding as to what zero pronouns stand for. He conducted an experiment with 30 Japanese 3-year-old children, in which the children were asked to identify zero pronouns. Based on this experiment, Nakayama reports that the children identified zero pronouns at the 82% accuracy.

[3] There are several morphological ways to focus on result. Examples include the following.

(i) a. Neko-ga sinde-iru "A cat is dead [a cat is in the state of having died]"
cat-Nom die
(cf. A cat is dying.)

b. Kisha-ga tuite-iru "A train is here [a train is in the state of having arrived]"
train-Nom arrive
(cf. The train is arriving.)

(ii) a. Namae-ga kaite-aru "The name is written [the name is written as a result of someone having written it]"
name-Nom write

b. Biiru-ga katte-aru "Beer has been bought [beer is here as a result of someone having bought it]"
beer-Nom buy

Ikegami (1985) also gives the following example and similar to it to illustrate a wide spread phenomenon of focusing result (p.275).

(iii) Wakasitakedo wakanakatta. "*I boiled water, but it didn't boil."

[4] Three pairs of intransitive and transitive verbs emerged at the same time.

[5] In his examination of two Japanese children's speech between 22 and 24 months of age, Rispoli (1987) finds that these children are able to distinguish between transitive verbs and intransitive verbs in general, and attributes the children's clear sense of transitivity to the semantic and pragmatic notions of animacy of the locus of change and the planned nature of an action. He further comments, "[i]ntransitive sentences were always more frequent when the theme referent was a true animate, animate surrogate, or vehicle… For both Jin and Hiroki, if a verb was produced in a sentence that referred to a planned action, that sentence was almost always transitive…" (p.195) Rispoli's analysis is not totally inconsistent with what I have claimed, but the second part of this quote is puzzling in

light of morphologically related transitive and intransitive verb pairs that we have been considering here. Recall that the event that is denoted by the verb pairs under consideration is change of state and change of location/position, and fits Slobin's (1985) characterization of the prototypical transitive event. That is, the event expressed by a transitive verb is consistent with Rispoli's description of transitive sentences, i.e., a planned action. If his reasoning were also to serve as a foundation for initial learning of a verb, then it would be strange that Aki's type count shows a relative ease in initial learning of intransitive verbs compared to their transitive counterparts: that is, a prediction should be reverse of Aki's situation.

7. REFERENCES

Bowerman, Melissa. "The Origins of Children's Spatial Semantic Categories: Cognitive Versus Linguistic Determinants." Gumperz, John J. and Stephen C. Levinson (eds.) *Rethinking Linguistic Relativity*. Cambridge: Cambridge University Press, 145-176, 1996a.

Bowerman, Melissa. "Learning How to Structure Space for Language: A Crosslinguistic Perspective." Bloom, Paul, Mary A. Peterson, Lynn Nadel, and Merrill F. Garrett (eds.) *Language and Space*. Cambridge: MIT Press, 385-436, 1996b.

Choi, Sonja. "Language-specific Input and Early Semantic Development: Evidence from Children Learning Korean." Slobin, Dan (ed.) *The Crosslinguistic Study of Language Acquisition, vol. 5: Expanding the Contexts*. Hillsdale: Erlbaum, 41-133, 1977.

Choi, Sonja and Melissa Bowerman. "Learning to Express Motion Events in English and Korean: The Influence of Language Specific Lexicalization Patterns." Levin, Beth and Stephen Pinker (eds.) *Lexical and Conceptual Semantics*. Oxford: Blackwell, 83-121, 1991.

Gentner, Dedra. "Why Nouns are Learned Before Verbs: Linguistic Relativity Versus Natural Partitioning." Kuczaj, Stan A.II (ed.) *Language Development* vol. 2. Hillsdale: Erlbaum, 301-334, 1982.

Gentner, Dedra and Lera Boroditsky. "Individuation, Relativity, and Early Word Learning." Bowerman, Melissa and Stephen Levinson (eds.) *Language Acquisition and Conceptual Development*. Cambridge: Cambridge University Press, 215-256, 2001.

Ikegami, Yoshihiko. *"Suru" to "Naru" no Gengogaku* [Linguistics of "Do" and "Become"] Tokyo: Taishukan, 1981.

Ikegami, Yoshihiko. "'Activity'-'Accomplishment'-'aAhievement'—A Language That Can't Say 'I Burned It, But It Didn't Burn' and One That Can." Makkai, Adam and Alan K. Melby (eds.) *Linguistics and Philosophy: Essays in Honor of Rulon S. Wells*. John Benjamins, 265-304, 1985.

Jacobsen, Wesley. *The Transitive Structure of Events in Japanese*,. Tokyo: Kuroshio, 1992.

Kageyama, Taro. *Dooshi Imiron* [Verb Meaning]. Tokyo: Kuroshio, 1996.

MacWhiney, Brian. "The CHILDES Project: Tools for Analyzing Talk," 2nd ed. Hillsdale: Erlbaum, 1995.

Miyata, Susanne. "Wh-Questions of the Third Kind: The Strange Use of Wa-Questions in Japanese Children." *Bulletin of Aichi Shukutoku Junior College*, No. 31: 151-155, 1992.

Miyata, Susanne. "Japanische Kinderfragen: Zum Erwerb von Form – Inhalt – Funktion von Frageausdruecken." Hamburg (OAS), 1993.

Miyata, Susanne. "The Aki cCrpus – Longitudial Speech Data of a Japanese Boy Aged 1.6-2.12." *Bulletin of Aichi Shukutoku Junior College*, No.34: 183-191, 1995.

Nakayama, Mineharu. *Acquisition of Japanese Empty Categories*. Tokyo: Kuroshio, 1996.

Nomura, Masami and Yasuhiro Shirai. "Overextension of Intransitive Verbs in the Acquisition of Japanese." *Child Language Research Forum* 28: 233-242, 1997.

Oshima-Takane, Yuriko and Brian MacWhinney. (eds.) CHILDES Manual for Japanese, 2nd ed., McGill University/Chukyo University, 1998.

Rispoli, Matthew. "The Acquisition of the Transitive and Intransitive Action verb Categories in Japanese." *First Language* 7: 183-200, 1987.

Rispoli, Matthew. "The Acquisition of Verb Subcategorization in a Functionalist Framework." *First Language* 11: 41-63, 1991.

Shibatani, Masayoshi. *The Language of Japan*. Cambridge: Cambridge University Press.

Slobin, Dan. "Crosslinguistic Evidence for the Language-making Capacity." Slobin, Dan. (ed.) *The Crosslinguistic Study of Language Acquisition, vol. 2: Theoretical issues*. Hillsdale: Erlbaum, 1157-1249, 1985.

Slobin, Dan. "Learning to Think for Speaking: Native Language, Cognition, and Rhetorical style." *Pragmatics* 1.1:7-25, 1991.

Smith, Linda and Maria D. Sera. "A Developmental Analysis of the Polar Structure of Dimensions." *Cognitive Psychology* 24: 99-142, 1992.

Talmy, Leonard. "Lexicalization Patterns: Semantic Structure in Lexical Forms." Shopen, Timothy. (ed.) *Language Typology and Semantic Description*, vol. 3, Cambridge: Cambridge University Press, 36-149, 1985.

Teramura, Hideo. "'Naru' Hyoogen to 'Suru' Hyoogen: Nichiei 'Tai' Hyoogen no Hikaku" ['Become' expressions and 'do' expression: comparison in 'voice' expression between English and Japanese]. *Nihongo to Nihongo Kyooiku Mojihyoogenhen*. National Language Center, 49-68, 1976.

Tomasello, Michael. *First Verbs*. Cambridge: Cambridge University Press, 1992.

Tomasello, Michael. "Pragmatic Contexts for Early Verb Learning." Tomasello, Michael. and William Merriman (eds.) *Beyond Names for Things*. Hillsdale: Erlbaum, 115-146, 1995.

Tsujimura, Natsuko. "Japanese Enter/Exit Verbs Revisited: A Reply to Kita (1999)." *Studies in Language* 26: 165-180, 2002.

Part 3

Stages in the development of verb grammar and the role of semantic bootstrapping

ELLEN HERR-ISRAEL AND LORRAINE MCCUNE

DYNAMIC EVENT WORDS, MOTION EVENTS AND THE TRANSITION TO VERB MEANINGS

Abstract. The purpose of this research was to examine evidence for the proposal that dynamic event expressions (relational words in English) in the single word period are based on sensorimotor cognition and the development of mental representation, and serve as a cognitive/semantic bootstrap to verb meaning and initial word combinations. Motion event semantics provided a framework for consideration of single dynamic event expressions in relation to first verbs used in sentences. Three English-speaking children were video-recorded monthly from 14 to 24 months, and their use of predicates expressing dynamic action in single and multi-word utterances (both relational words and verbs) was examined. Findings were that: first, relational word use precedes the generalized use of any verbs; second, when verbs are initially used, primal verbs, defined both by their universal early occurrence and their relation to the motion event situation, dominate; and third, combinations that include a relational word, a primal verb or both constitute the majority of the earliest combinations. These results support the proposal that dynamic event expressions serve as a bootstrap in the transition to verb use and word combinations.

1. INTRODUCTION

R. Brown, in opening the modern field of empirical child language study, noted that, "it is my impression that the first meanings are an extension of the kind of intelligence that Jean Piaget calls sensorimotor" (Brown 1973:198). This hypothesis fell into disfavor as studies showed only weak correlations between sensorimotor and language milestones. However, work relating the cognitive semantics of single words to sensorimotor understandings (McCune-Nicolich 1981), and more recently to motion event analyses (McCune, in press; Talmy 1975, 2000; Vihman 1999) suggest that such early cognitive underpinnings to linguistic development may actually exist, although in much more complex form than was originally anticipated.

The proportionally low frequency of verb use in the single word period in many languages, including English is well documented (Bornstein et al., 2004). Instead, during the months when children are limited to one word at a time, their lexicon is primarily composed of referential nominals and dynamic event words.[1] Nominals refer to moveable and manipulable objects such as toys, food, or body parts and animate beings such as family members or pets. Dynamic event words make reference to dynamic states and movements, rather than entities. (e.g., *up*, *more*, *allgone*, McCune-Nicolich 1981). In this

N. Gagarina and I. Gülzow (eds.), The Acquisition of Verbs and their Grammar: The Effect of Particular Languages, 125–149.

chapter we propose that dynamic event words have their basis in: a) children's sensorimotor knowledge of the effects of their own actions, b) motion event understanding, and c) mental representation, allowing initial cognition and verbalization of dynamic event meanings shaped in accordance with the resources of the ambient language.

Verbs are virtually absent during the single word period in languages such as English, occurring only infrequently accompanying child action. Ingram et al. (this volume) augment this assertion in their discussion of verb use in the 'holophrase' period. Critical spatial and temporal relations in English are expressed with particles, which are easily segmented and occur as single relational words. Consequently the dynamic aspect of events in English is, at first, expressed primarily or exclusively by relational words. In Korean and similar languages, in contrast, the same sorts of early dynamic event meanings are expressed using a special set of spatial verbs (Bowerman and Choi 2001). In the early word combining period dynamic event words in English comprise both relational words (also occurring in the single word period) and all verbs used.

We propose that the learning and use of relational words, or dynamic event words including spatial verbs, provides a cognitive/semantic bootstrap to initial, or (in languages like Korean) more general verb meanings. Use of language to reference spatial and temporal aspects of events provides transitional cognitive and linguistic experience. These experiences, under conditions of exposure to verb use in a native language, yield development of or expansion of verb meanings. During this same period of single word use, children's continued cognitive development, both in the relational understanding of space and time and in mental representation, contributes to the onset of dynamic event words including first verbs. Vihman (1999) demonstrated a developmental pattern in her bilingual son (exposed to Estonian and English) supporting this proposal.

Ninio noted "a rather unexamined vague consensus that the most 'natural' or prototypical semantics underlying transitive constructions is one of a volitional animate actor strongly affecting the state of an inanimate patient by some overt dramatic action" (Ninio 1999a:623). Bowerman found that, rather than demonstrating such an innate thematic basis for linking verbs with their arguments as suggested by Hopper and Thompson (1980) and Pinker (1984), "acquisition of verbs of different kinds and the accuracy with which their arguments are mapped is inconsistent with what we should expect under the assumption that knowledge of linking is innate" (Bowerman 1990:1285). In fact, Bowerman's research also failed to support early learning of such rules, although evidence for their existence and generality in children of seven to nine years of age suggested rule learning was operative at that time.

However, there is cross-linguistic evidence that at the beginning of verb use children tend to use a set of verbs that share meaning with earlier dynamic event words, as we will show, while adding semantic elements that complement dynamic event word meaning. Ninio found that the earliest transitive verbs, rather than involving highly active agents, are general in nature (e.g., *put, make*) and "express fundamental … relations of incorporation into, and ejection from the personal" (Ninio 1999a:619). Similarly, Martinot (this volume) notes that the most frequent early verbs used by the French-learning children in her study were "verbs with very light meanings." The prototypical concept underlying initial syntactic transitivity is, according to Ninio, the child's physical relationship with objects. This finding echoes work by Clark (1978), who noted the prominence of *do* and *make* in reports of early verb use, as did Bowerman (1978), remarking the importance of general action verbs such as *put*, *give*, and *make*. In her 1990 report Bowerman provides more complete documentation of the earliest verbs, results conforming to the Ninio findings. Tomasello (1992) included relational words referencing dynamic events in his designation of first verbs, and his daughter (T), the subject of his own investigation contributed data to the Ninio study. Despite variation in interpretation, these previous researchers have identified similar sets of 'first' verbs.

In our view children's established capacity for expressing dynamic event meaning, along with dawning recognition of the self as agent and experiencer provides the basis for the initial development of verbs. The particular verbs most commonly identified as "first" (e.g., Bowerman 1990; Clark 1978; Ninio 1999a; 1999b) support this claim. Recent analyses (McCune, in press; McCune and Vihman 1999; Vihman 1999) demonstrate that meanings expressed in relational words referencing dynamic events reveal the salience of motion event understanding. This same underlying structure is proposed as the basis for first verbs. The meanings of children's dynamic event expressions also demonstrate their spatial and temporal understanding of the events referenced and entail sensorimotor knowledge and a capacity for internal mental representation. In the following sections we (1) summarize the cognitive bases of language in the early period, (2) analyze the manner in which English relational words express the space/time aspects of the motion event situation and (3) review prior data on early dynamic event reference including relational words and early verbs. A study of the transition to verb use in three children learning English is then reported.

2. MENTAL REPRESENTATION AND LANGUAGE ACQUISITION

According to Werner and Kaplan (1963), children's initial limitation to single word expression reflects a symbolic capacity that is similarly limited to global representation of events. The single word marks for the child the complex global notion of a salient entity as well as the environmental aspects affecting it. This does not mean that the single word actually represents a linguistically complete underlying sentence**.**

Rather, nominal words emphasize the entity or object named but include some sense of the object's participation in the event. A child who says *ball* is also noting the movement of that object as it rolls away. A child who says *allgone*, having finished her juice, includes some sense of the past experience of a full cup of juice, and the empty cup present in her hand as aspects of her conscious experience. From a representational perspective it is critical to note that the child's use of *allgone* in this example demonstrates conscious, that is representational, attention to brief moments of past and potential future, as well as to the present. Dynamic events occur extended in time effecting changes in the perceived present. Verbs referring to dynamic action events of all sorts also rely on this representational capacity, first expressed by children learning English with single relational words.

Children's actions with objects, including interest in creating configurations where one object occludes another and is then retrieved, and sequences where objects are repeatedly placed in a location in relation to the self (Sinclair et al. 1989) provide the basis in experience for developing meanings coded in the ambient language by dynamic event words. Observation of the movements of objects and persons also contributes to this representational knowledge.

McCune has demonstrated that the onset of global mental representation, allowing children to represent an event by a simple action (e.g., putting a spoon to the lips representing a meal), is demonstrable in children by the time they begin single word use (Nicolich 1977; McCune-Nicolich 1981; McCune 1995). When children are limited to single words they use expressions from the ambient language which focus either on entities in the environment or on dynamic aspects of events in which objects and persons participate. The parallel development of the two classes of words (entity words and dynamic event words) suggests that both aspects are salient to the child (even though each may refer to whole events in the one-word period). Different classes of single words focus on the entity or dynamic aspects of events that will later be referenced sententially using both dynamic event and entity words (Gentner 1978; Lyons 1968).

In representational play children move from pretending single acts to combining pretend acts. For example, instead of merely giving a doll a drink from a cup, they may first pretend to pour imaginary liquid into the cup then give the doll a drink. In language dynamic event word use in the single word period already testifies to the child's potential for complex combinatorial representation. Representational consideration of the previous or future situation – that is, perceptual experiences removed from the present is implied in dynamic event word use. (Recall the use of *allgone* above.) By the time children begin to use word combinations they have also demonstrated the ability to combine symbolic play acts (McCune 1995). This new more differentiated representational ability allows combinations including both a dynamic event word and an entity word (e.g., *more juice*).

Combinatorial representation is also essential to the use of typical early verbs. Children are unlikely to remark: *do* or *want,* as these very general early verbs cannot stand alone. Rather, once the capacity for representational combination is available we find, *I do* or *want juice*, as the child combines an agent or patient with the dynamic event expression (*do* or *want*). Two typical forms of word combinations testify to the child's changing representational ability. In one type, learned routines such as, for example, *I do it,* the reference to action remains somewhat undifferentiated and global, as the lexical elements do not occur independently. However, the inclusion of elements referring to self as agent and a generalized patient suggest beginning differentiation of the action event into agent, action, patient components.

In the other type of early combination words are juxtaposed, each with its own salience in relation to context as, for example, *daddy car*, perhaps said as his father enters the family car. In this case the child's focus shifts between elements that he clearly sees as related, without linguistically specifying their relationship. The combination of words from two different languages in bilingual children demonstrates children's potential for inventing such combinations (Vihman 1999).

In contrast with these earliest combinations, a subsequent period yields combinations that are based on initial ordering rules which may be idiosyncratic to the child, or conform to common sequences in the input, reflecting a more advanced symbolic capacity (McCune 1995). McCune-Nicolich and Bruskin (1981) found that such patterns occurred only after the onset of more sophisticated symbolic play in which play acts are planned prior to execution or where one object substitutes for another (Nicolich 1977; McCune 1995). We believe that the capacity for an internal mentally represented intention to direct external behavior accounts for this behavior in play. In language 'positional patterns,' where individual lexical items retain a statistically consistent position across two-word utterances may depend on

this same internal representational capacity. (Braine 1963, 1976; McCune-Nicolich and Bruskin 1981).

The existence of positional regularity, varying one element at a time, demonstrates the initial capacity to link linguistic symbols directly by an ordering principle, rather than through each word's independent reference to elements of context. A prior semantic intention to express a particular meaning is conveyed by the ordered relationship of the two words produced. This positional regularity is probably a temporary production device, easing linguistic expression during the period when the child is gradually learning the grammatical conventions of her ambient language. With advancing representational development and experience using language, word order and meaningful morphological variation come to reflect grammatical structure (Veneziano 2003). In addition, the frequency of exact deferred imitation, or the existence of "frozen phrases" (Lieven, Pine and Baldwin 1997) decreases.

3. MOTION EVENTS, SYMBOLIC DEVELOPMENT, AND LANGUAGE

From birth onward, experience with entities in motion facilitates sensorimotor understanding of the properties of Time and Space, which will later be represented mentally and coded with language (Piaget 1954; Sinclair et al. 1989). First, motion event understanding is limited to perceptual "taking in" of events that are present and physical construction of spatial relationships by purposefully placing objects or by ones own movements. Gradually the child becomes able to mentally represent – that is remember and predict – the elements of an event sequence from conditions at a given time point in the sequence. For example, a child who places an object in a container, hiding it from view, then subsequently retrieves the object, must keep its location and potential movements in mind during the brief moments when the object is not perceptually present. Saying *allgone* while the object is occluded provides linguistic evidence of the child's representation of the entire sequence. Observation and enactment of such movements of entities and people in space over short durations of time form the basis upon which symbolic ability and, later, language are built. This sensorimotor and early representational understanding can be interpreted as motion event understanding.

Talmy (1975, 2000) proposed that humans' universal common experience and conceptualization of motion events might form the basis of later linguistic relationships, and provided semantic analyses conceived by parsing sentences into common semantic constituents of motion. A motion event is considered a "situation containing movement of an entity or maintenance of an entity at a static location" (Talmy 1985:60). Movement is defined to include directed or translative motion that results in a change of location, and "location" can

include both static location or contained movement that results in no overall change in location, such as jumping up and down (Talmy 1985:60; Choi and Bowerman 1991:85). The semantic components characterizing a motion event, which are expressed sententially in adult speech, include movement, figure (the moved object), ground (with respect to which movement occurs), path (the direction of motion in relation to the ground), deixis (direction with respect to the speaker), manner (e.g. *walked, ran, rolled* vs. *moved)* and cause (e.g. transitive vs. intransitive; Talmy 1975).

Returning to the issue of the child's developing symbolic capacity, one can say that the most basic differentiation within the "whole" of what one can communicate about an entity (predication) is how the entity relates to other aspects of the motion event – i.e. whether it is moving or stationary, its location with respect to ground and the speaker, whether it is visibly present or occluded, and where it exists at a given moment in time. Because the human perceptual system is biased toward motion (as is that of other animals) the most salient aspect of entities, other than their nature as such, is their motion in space over time. Developmentally earlier use of dynamic event words rather than, for example, color or size terms indicates that object motion is linguistically differentiated from the object itself prior to differentiation of such inherent characteristics.

Motion of an object renders the notions of *here* and *there* and *before* and *now* visible and therefore tangible. For example, before a child has the capacity for mental representation, she experiences space and time from her vantage point in the present and through the motion of objects both animate and inanimate. The infant in his infant seat experiences space and time by watching the movement of his mother in the room. He perceives her as smaller as she walks away, which he learns to mean that she is at a distance from him. He may experience this distance with distress – he feels the difference between now when she is distant and *before* when she was close. As the child develops motorically and cognitively he is able to experience space and time through his own movement of objects within his space. These perceptual and motor experiences facilitate the child's ability to mentally represent the structure of space and time via motion events as he experiments with motion events and relational meanings. To put it succinctly, interacting with objects while observing their motion and location aids the understanding of time and space by helping to make the intangible notions of time and space more tangible (Sinclair et al. 1989).

McCune and Vihman (1999) analyzed the basic set of meanings described by McCune-Nicolich (1981) and proposed three super-ordinate categories of dynamic event expressions reflecting the semantics of motion events (a subset of the elements of the motion situation as described by Talmy 1975, 2000).

These categories are: 1) **Spatial Direction** or Path in the vertical plane (e.g., *up*, *down*) or the "deictic plane", involving movement toward or away from the child's body (e.g., *here* used in exchange; *there* accompanying placing actions); 2) **Spatial Relations between Entities** (Figure/Ground), including reversible aspects of containment (*open, close, in*, *out*), attachment (*stuck*) and occlusion (*allgone*); and 3) **Temporal Event Sequences**, where the relational word indicates a mental comparison of the ongoing state of affairs to a prior, expected, or desired reverse alternative: (a) iteration or conjunction (*more, again*) and (b) potential reversibility or negation (*no, uhoh, back*).

4. DYNAMIC EVENT WORD USE

Cross-linguistic studies have demonstrated that the world of single dynamic event words is carved up differently across languages (e.g. Bowerman 1994; Bowerman and Choi 2001; Choi and Bowerman 1991; Vihman 1976, 1985, 1999; Leopold 1939). However, these same papers demonstrate that the aspects of events and situations that children choose to express with relational words in English (or, in languages like Korean, with spatial verbs) is common across languages (McCune, in press). Bowerman and Choi (2001) note that children's first spatial morphemes "typically revolve around notions of containment, support, attachment, motion up and down the vertical axis, and opening and closing" (p.480). These authors demonstrate that generalization of these words respects their extension in the input language, with interesting variation across Dutch, Korean, and English.

English, German and Estonian data demonstrate the convergence of children's relational words on a common set of motion related meanings (See Table 1, adapted from McCune and Vihman 1997). Participants 1-3 were previously studied by McCune and provide data for the study of the transition to verbs, which is the subject of this report (Sandra (1), McCune-Nicolich 1981; Aurie (2) and Alice (3), McCune 1995) Participants 4 and 5 were studied by Vihman (Virve (4) 1976; Raivo (5) 1985; 1999) and Hildegard (6) by Leopold (1939). It can be seen in Table 1 that all of the children produced words in the superordinate categories (Spatial Direction, Spatial Relations between Entities, and Time: Reversible Event Sequences (McCune and Vihman 1997, 1999). Three of these children were bilingual speakers of English and another language (Estonian or German). In each case relational words in English are noted in the first column of the child's data). Considering all of the subordinate categories (Vertical Path, Deictic Path, Containment, Attachment, Occlusion, Iteration/Conjunction and Negation/ Reversal), all six children also produced words in each subordinate category, with the exception of Attachment (N=4). This consistency in early dynamic

event word use across children learning different languages supports a perceptual/cognitive interpretation for the categories proposed. That is, perceptual/cognitive background knowledge is shaped by the resources of the ambient language to yield a highly similar initial motion semantics which nonetheless respects the categories of the adult language in its application.

The children studied by McCune-Nicolich (1981; 1995), Vihman (1976, 1985) and Leopold (1939) produced these dynamic event words prior to generalized use of true verbs. If verb forms were used in the early period they were restricted to specific event situations, usually accompanying child action, or to carrying the same kinds of meaning as relational words.

5. FIRST VERBS

Ninio (1999a) examined the longitudinal language development of 16 children – one learning English, the others learning Hebrew. She demonstrated that the first verbs children use in combination are relatively small in number and are semantically highly general. Ninio noted that first transitive verbs fall into three general categories: 'obtaining' verbs (*want, get, give, take, bring, find)*; 'creating' verbs (*do, make, prepare, build, draw);* and 'consumption/perception' verbs (*eat, drink, see, hear*).

Results indicated that there was a significant overlap in the earliest verbs of the different children. Ninio analyzed the first two verbs used in verb/object (VO) combinations and the first two that were used in subject/verb/object (SVO) combinations. With regard to the verbs used first in VO combinations, there were only 11 verb types used across the 16 children. The majority of these (78%) were 'obtaining' verbs (*want, take, give*). The set of first verbs used in SVO combinations was similar to the VO set but had a different composition in terms of proportions of verbs from each of the three categories of meaning ('obtaining', 'creating', 'consumption'). There were 13 verb types in the set, 37% were obtaining verbs, 33% were creating verbs (*make/do, build*), and 11% were consumption verbs (*eat, drink,*). Ninio considered these verbs as "pathbreaking", because they tend to be a harbinger of rapid subsequent verb learning.[2] It is as if these verbs allow the child to 'break into' the category *verb* and thus facilitate the transition to VO and SVO combinations.

In a complementary report Ninio (1999b) found that first intransitive verbs for 20 Hebrew-learning children were *come, go,* and *fall.* An analysis of verb production by 48 mothers of an independent group of early language-learners confirmed that these three verbs were also most frequent in maternal speech. In reviewing the semantic roles for both transitive and intransitive verbs Ninio (1999a andb) documents that these highly frequent and

early-learned verbs encode meanings most likely to be grammaticalized across the worlds languages. McCune (in press) proposed that the cognitive significance of these words is also implicated in the tendency for them to emerge as grammaticalized elements.

We consider the kinds of first-occurring, general meaning verbs that Ninio and other researchers have identified as "primal verbs". By this designation we mean to imply both their temporal primacy and the foundational nature of their meanings. In our view the path-breaking effects of these verbs follows from more fundamental characteristics: first, their basis in the dynamic event meanings established in the single word period, and second, their simple expression of agency applicable across a variety of general action scenarios.

6. DYNAMIC EVENT WORDS AS A BOOTSTRAP TO VERB LEARNING

Why do children, when they are limited to single words, reference only a limited set of dynamic events rather than developing a broad repertoire of verbs to convey more varied dynamic event meanings? Why do children, in the early word combining period, continue to rely on dynamic event words, primal verbs or both to convey dynamic event meanings? Is there a developmental link between early dynamic event expressions and use of primal verbs?

English-learning children's relational words express dynamic events and are inherently predicative. They provide experience with communicating meanings that are partially representational (rather than relying strictly on perceptual experience). Unlike verbs, they express a single, fairly simple aspect of motion in space occurring over a short duration of time. We propose that some of these same meanings characterize the primal verbs. The transition to verbs is further fueled by increasing capacity for representation and the growing understanding of self as agent and experiencer. The group of primal verbs identified by Ninio (and others) reflect these processes.

The proposal that dynamic event word use in the single word period serves to bootstrap word combinations with primal verbs follows from the semantic and developmental connection between English-language relational words and primal verbs. (In other languages analogous dynamic event meanings expressed in the single word period are hypothesized to have similar effects). Relational words are inherently predicative because their meaning specifically expresses temporal or spatial relationships between entities, between self and entity, between self and other or between entity and other. With regard to the motion situation, the focus of relational words is the object's relationship to its surroundings (see Table 1) Deictic path words

(e.g., here, mine) emphasize location in relation to the child speaker, and are sometimes used in exchange routines, demonstrating the notion of spatial reversibility. Vertical path expressions identify location or trajectory in the vertical plane (e.g., up). Relational words expressing figure/ground relationships focus on the spatial relationship between a target (the figure) in relation to another object or location (the ground; e.g., *stuck*, while trying to remove a puzzle piece, or *allgone*, as a toy falls from the high chair tray.)

Often these utterances imply the end-state, the static relationship of an object with its environment. For example, when a child who has not yet begun to combine words says *more* (when she wants another cookie), her focus is on the imagined end state, which is having another cookie.

As the child gains more proficient mobility, more physical distance from her mother, more experience with her own agency and the presence or absence of agency of the other, she becomes increasingly aware of self/other distinctions and the notion of her own and others' agency. Now children have the representational ability and the desire to talk about their own role in movement of objects or the role of another person in these dynamics. Thus many primal verbs focus on agency in the sense of moving objects to and from the self and to and from the other – that is they are primarily deictic path verbs. Continuing the example used above, with regard to a child wanting another cookie she may now say *want cookie.* Here the emphasis is on her **desire to obtain** the cookie, that is, to bring the cookie **closer** to the self and into her own domain of control, rather than the iterative end state implied by *more*.

Table 1. Dynamic Event Meanings Expressed by Children Learning 3 Languages

Relational Words in Motion Categories	*Participants*[3]								
	1	2	3	4E	4S	5E	5S	6E	6G
Spatial Direction									
Path: Vertical									
Down	X	X	X	X	X		X	X	
Up		X	X	X	X	X	X	X	
Path: Deictic									
here/thanks	X	X	X		X		X	X	X
Mine		X		X					
There	X	X	X	X		X	X		
Spatial: Figure/Ground									
Containment									
Open	X	X	X	X	X		X	X	
Closed					X		X	X	X
Out		X	X	X	X		X	X	X
In			X	X	X		X	X	X
Attachment									
stuck/fitted	X	X	X	X	X		X		
Unstuck[4]				X	X			X	X
Time: Reversible Event Sequence									
Occlusion/Absence									
Allgone		X	X	X	X		X	X	X
'bye	X	X	X	X					
Peekaboo				X					
Iteration/Conj.									
More	X	X	X	X	X		X	X	
Again				X	X				
Negation (Reversal)									
No	X	X	X	X			X	X	
Uhoh	X	X		X					
back	X	X					X		

According to our analysis, primal verbs can be shown to draw on the semantics of relational words in English while incorporating the notion of agency and offering more nuanced details of meaning. Nino (1999a) proposed that the prototypical concept underlying syntactic transitivity is not a situation

of an agent that is highly affecting a recipient of the agent's action. Instead, the prototypical concept underlying syntactic transitivity is, according to Ninio, the child's physical relationship with objects. Recall that Ninio categorized the set of first verbs used as 'obtaining', 'consumption', and 'creating' verbs, and stated that they deal with the "literal and metaphoric ingestion and ejection of an object by a human being" (Ninio 1999a:644). These notions are the essence of deictic path toward or from the self.

The 'obtaining' verbs depict a person taking temporary possession of an object, or bringing it within the personal domain. The 'consumption' verbs depict a person physically ingesting an object, and the 'perception' verbs depict a kind of mental ingestion. Both of these meanings express deictic path toward the self. Finally, the 'creation' verbs "depict the physical or metaphorical discharge of an object from the body or person of a human being" (Ninio 1999a:644-645). Many of Ninio's 'creating' verbs are verbs that express the motions of agency (usually of the self) in a nonspecific way (e.g., *make, do, play*, *can*). Metaphorically this can be interpreted as deictic path from inside to outside the self. Indeed the notion of "agency" or "motive to act" is an internal experience that finds expression in external action.

Primal verbs such as *find* (included by Ninio in the 'obtaining' category), *close,* and *hide* express figure/ground relationships. The more general figure/ground relational word, *allgone,* expresses the absence of some entity to perception, or its absence in a given location. The primal verbs noted include aspects of the *allgone* notion, but present nuanced meanings and include attention to an agent who effects some change in the occlusion situation, or use a specific verb designating retrieval of the missing object by a presumed agent (e.g., *find*).

The purpose of the current research is to examine evidence for the proposal that English language relational word use in the single word period serves as a bootstrap to verb meaning and word combinations. Evidence in support of this proposal would include: 1) relational word use precedes the generalized use of any verbs, 2) when verbs are initially used, primal verbs defined in relation to the motion event situation dominate and, 3) combinations that include a relational word, a primal verb or both should constitute the majority of the earliest combinations used.

7. METHOD

Participants were three children (all girls) whose monthly play sessions with their mothers were videotaped and transcribed according to the procedure described below. Two of the children (Aurie and Alice) were originally part of a study of symbolic play and language (McCune 1995). The other child

(Sandra) was part of a study of relational word use in the single word period (McCune-Nicolich 1981). These play sessions were conducted longitudinally beginning when Aurie and Alice were 9 months of age and when Sandra was 10 months of age. The studies continued until the children were 24 months of age. Data in the present study comprise the speech records of the children from 14 to 24 months of age. The mother/child dyads were sought via newspaper advertisements and word of mouth. Social class was not a criterion for selection, but all parents were middle class on the basis of their education, employment, and area of residence. In all cases mothers were primary caregivers for their children, filling that role for 50% or more of the child's waking hours. English was the only language spoken in the homes.

7.1. Procedure

Data were collected in homes at a time of day when mothers indicated that their children were usually awake and engaged in play activities. For the play observation, mother and baby were seated on the floor in a room free of the child's own toys; the set of toys similar to those used by McCune (1995) was placed near them. This large diverse toy set is designed to offer manipulative, problem-solving, and pretend play opportunities. The toys were arranged in a large plastic bucket, with the same toys protruding at each session. This method of presentation typically leads to exploration of the contents of the bucket, which provides both familiarization with the toy set and a warm-up exploratory period before the child settles into play. The play sessions were videotaped using an external microphone placed near the mother and child.

7.2. Transcription and Language Evaluation

Transcriptions of children's language were made with accompanying contextual descriptions of the child's actions, the mother's actions and language. *Words* were identified by the following criteria: proximity to adult phonological shape, appropriate context, and multiple occurrences. Each transcription was checked by an independent transcriber, and cases in which disagreements concerning the status of an utterance as a word could not be resolved were eliminated from the analysis. Children were considered to have begun lexical development if they produced five or more spontaneous single words. Vihman and Miller (1988), in weekly observations, found that subsequent to, but not prior to this milestone, children showed stable vocabulary acquisition.

Word combinations were identified as sequences of words sharing a single intonation contour with falling intonation only on the final word. Vocalizations that are usually considered unitary expressions in the early

language literature (e.g. *what's that*, *thank you*, and *allgone*) were treated as single words both because they always function as a unit and because they are frequently observed when all of the child's other language consists of single words. Onset of multiword utterances was defined as production of five or more multiword types in a given session to correct for the possibility that early combinations may reflect idiosyncratic errors of segmentation. Mean length of utterance (MLU) was computed in words, rather than morphemes. This is a conservative approach, customary at the early transition to combinations because data are insufficient to determine the independent status of a particular morpheme. For example, the plural "s" may be observed with only a single lexical item (e.g., *shoes*) for which the singular form may fail to be observed.

Utterances that were direct imitations of all or part of the mother's previous utterance and those that were considered to be routinized expressions (recitations of children's poems or songs, e.g., *on the treetop* and compound-like combinations, e.g., *hurry up*, *garbage can*, *honey suckle*) were not included in the analysis.

The criteria for identifying relational words were as follows: 1) the word occurred persistently (at least 3 times) and was not an imitation of what the mother had just said, and 2) the word made specific reference to real or potential states and transformation of objects, people and events over time.

The relational words identified in McCune-Nicolich (1981) were: *up, down, in, out, uhoh, oops, open, close, here, there, more, back, allgone, stuck, no* (when used to express potential reversibility or negation in the sense that the child prefers a reversed state of affairs), *bye-bye* (if used to express the disappearance of something or under conditions of presence when the child is about to cause a disappearance), and *okay* (when it is accompanied by the child's action, to mark the completion of an action).

The procedure for identifying primal verbs was to first identify all verbs used by each of the three children. From this set of words we excluded those that occurred only as imitations or in routinized phrases and forms of the verb 'to be'. Next, we evaluated each verb with the criteria that it: 1) have very general, content-neutral meaning, 2) that it express some aspect(s) of the motion situation, and 3) that the verb is used by at least two of the children and in two different combination types by each child. Verbs used that do not conform to these criteria, non-primal verbs, were considered as a distinct group and are described more fully in the results section. Table 2 provides an overview of the children's speech.

Table 2.

	Onset of Combinations (at least 5 in a session)	*Number of Multiword Utterance Types (Onset to 24 months)*[5]	*MLU at 24 months*
Aurie	20 months	240	3.26
Sandra	18 months	263	2.75
Alice[6]	17 months	1141	3.88

8. RESULTS

8.1. Single Word Period

We found that in the single word period the children almost exclusively used relational words rather than verbs to reference dynamic events. In addition to relational words identified by McCune-Nicolich (1981), we identified (during the total observation period) the following: *another, away, around, empty, outside, next, over, other, where, some, at, from* and *now. Empty* and *where* were used as single words – the others appeared only in combinations. Table 3 presents the proportions of each type of dynamic event word (relational word, primal verb or other verb) used by each child in the single word period. For this analysis, each dynamic event word was 'counted' in the month in which it first occurred and not thereafter (cumulative type frequency).

Use of relational words for expression of location or movement dominates in the single word period for all three children. The only dynamic event word that Aurie used in her single word period that was not a relational word was 'push' which she used when trying to get a puzzle piece into its correct place, expressing difficulty with a figure/ground relationship. One-third of Sandra's dynamic event words (4 of 12) in her single word period were primal verbs. The primal verbs she used were *eat, give, get,* and *go. Get, give* and *go* were used only in combinations (*get back, go in, I give it*) in her 18-month session (recall that the beginning of the word combining period was defined, for each child, as the session in which she had five or more different combinations, 20 months for Sandra). Alice used no primal verbs during her single word period (except *drink* which was imitated). The non-primal verbs she used were *clean* and *wipe*, both of which are verbs that are highly tied to the play actions she was producing at the time.

Table 3. Dynamic Action Words Used in the Single Word Period

	Total Single Word Types	*Dynamic Action Words as % of Total*	*Relational Words*	*Primal Verbs*	*Non-Primal Verbs*
			as % of Dynamic Action Words Used		
Aurie (16-19 mos.)	38	21% N=8	88% N=7	0	12% N=1
Sandra (16-19 mos.)	59	20% N=12	66% N=8	33% N=4	0
Alice (16-17 mos.)	55	11% N=6	67% N=4	0	33% N=2

8.2. Word-Combining Period

In the word combining period the children relied on both relational words and primal verbs to express location and motion (Table 4). The path breaking verbs identified in the Ninio studies (here considered as primal verbs) were: *bring, can, close, come, do, drink, eat, fall, find, get, give, go, hold, look, make, open, put, see, take* and *want*. In this analysis, verbs that were not specified in Ninio (1999a or 1999b) but that convey similar general meaning about movement and spatial relations with regard to the self are included in the primal verb category (e.g. *feel,* and *need*).

Multiword combination types that included a relational word and no other type of dynamic event word were on average 26% of total combination types. Combinations that included a primal verb and no other type of dynamic event word were, on average, 40% of combination types. Word combinations that included both a relational word and a primal verb, were on average 19% (Table 4, Column 4). So, when we combine the averages we find that 85% of the children's combinations included a relational word, a primal verb, or both.

Table 4. Dynamic Event Word Combinations During Word Combining Period

	Total Types	*% Including*			*Total % accounted for by relational words and primal verbs*
		relational word	*primal verb*	*both relational word and primal verb*	
Aurie (20-24 mos.)	202	28% N=57	51% N=103	9% N=19	88% N=178
Sandra (20-24 mos.)	250	31% N=76	33% N=83	19% N=48	83% N=208
Alice[7] (18-20 mos.)	320	20% N=64	35% N=112	29% N=93	84% N=269

8.3. Verb Meanings Observed

The three children used a total of 29 primal verbs (Table 5), accounting (along with relational words) for 85% of multiword types (Table 4), while 49 additional verbs, not meeting criteria as primal verbs accounted for less than 15%, of multiword types. Those combinations omitting any dynamic action word (e.g., *daddy car*) are also included in this small percentage. The primal verbs used fall into three categories pertinent to the motion situation. The majority of the verbs indicate path in the deictic plane, overlapping in meaning with relational words expressing Spatial Relations: Deictic Path. They more specifically "deal with the literal and metaphoric ingestion and ejection of an object by a human being" (Ninio 1999a:644). The second category of meaning is Figure/Ground relationships, overlapping in meaning with the relational word category Figure/Ground. Finally, the third meaning expressed is Agency in Motion, sharing some meaning with the relational word category, Temporal Event Sequence, but emphasizing the sense of agency.

These primal verbs are nonspecific in the sense that they do not express details of the action such as the manner of motion, and they are broadly applicable. However, unlike relational words, they do demonstrate the child's new knowledge of agency, essential to the use of verbs. Recall that Ninio identified three categories for 'generic transitivity verbs': 'obtaining',

'consumption' and 'creating' verbs. The verbs that Ninio placed in the 'obtaining' and 'consumption' categories are those that express Deictic Path. *Come* and *go* can also express deictic relationships. Ninio's 'creativity' verbs conform to our Agency in Motion. Some of our Figure/Ground verbs were also identified by Ninio, and considered 'obtaining' verbs.

The verbs in Table 5 are those identified as primal verbs according to the criteria cited earlier. Recall that one aspect of the primal verb definition is that it is used by at least two children and in at least two combination types by each of the two children. The verbs with an asterisk in table 5 are those that were used by all three children. Those in bold were also identified by Ninio. All 3 of the children used 18 of the 29 primal verbs identified across participants, and 2 children used the remaining 11 primal verbs. The average number of verb types used by the three children during the period studied was 33. So, 55% of verb types used were primal verbs used by all three children. Only one of the verb types in the nonprimal group, *sleep,* was used by all three children. It is noteworthy that the set of early verbs used by participating children in several chapters of this volume include many of the same verbs listed in Table 5. For example, see Tsujimura examples 10, 12 and 15; Martinot, figure 1 and Ingram et al., table 6 and example 3 (Dorothy's emerging verb classes).

Several verbs that the children used expressed meanings similar to primal verbs but were not included as primal verbs because they were not used by at least two of the children in two utterance types by each child. These verbs express either figure/ground (*spread, wear, wrap, fill, turn*), position or change of state (*bend, fix, lay, tie, untie, sit*), or deictic (*kiss, smell, touch, hate*) meanings. All of these 15 primal-like verbs were used by only one child except *sit* and *wear* which were used by two of the children. Interestingly, these less commonly used verbs, while similar to primal verbs in meaning, seem to be more semantically delimited (e.g., only certain entities can be *spread*, *bent*, or *tied*). The essence of primal verbs is that they do not have to incorporate additional semantics, but can be freely used across a broad range of situations. These verbs are probably also highly frequent in the input, so both simplicity of semantics and frequency of exposure would contribute to their early use).

Of the 34 non-primal verbs, 27 were used by only 1 child, 6 were used by two children (*read, call, sweep, wipe, talk,* and *clean*), and one (*sleep)* was used by all three children. Many of these non-primal verbs were used during play situations that included a 'prop' that was related to the verb in question. For example, Sandra used *wipe* and Alice used *clean* when playing with a mop. Other verbs in this category were *call* or *talk* (when playing with a toy phone), *dump* (when playing with a toy dump truck), *sweep* (when playing

with a broom), *swim (*when playing with plastic fish), *sleep* (when putting a doll to bed) and *brush* or *comb* (when playing with a brush or comb). Some fit Hopper and Thompson's (1980) notion of 'high transitivity' – namely, those that express events in which the agent acts on the patient in a physically obvious way, e.g., *hit, bite, catch, tickle* and *cut.* Finally, some non-primal verbs were only used as part of formulaic phrases, for example, *try (try it)* and *let (let me).*

Table 5. Categories of primal verb meanings

Category of Meaning	*Primal Verbs Used*
Space: Deictic Path	come,* ***drink,***[8] ***eat,***[9]* *feed, feel,* ***give,*** *gave* ***get,*** go,* *have,* help,* ***hold,*** like,* ***look,*** *leave, need,* ***put,*** ***see,*** ***take,*** took,* ***want****
Space: Figure/Ground	**close,*** **find**, found,* hide,*
Agency in Motion	**can**, **do**, play,* **make***

9. DISCUSSION

Ninio's identification of "model" or "path-breaking" verbs was based on their occurrence as the first two verbs in either VO or SVO expressions, without prior consideration of meaning in their identification. Nevertheless, she found coherence across children in the kinds of meanings expressed in these first verbs. Our thesis is that common cognition and motion event understanding, previously expressed in relational words in English, provide the foundation for this consistency across children.

In our analysis we found that relational words were the dominant means of referencing dynamic events in the single word period. This is not surprising since relational words purely and simply express temporal or spatial relationships between entities, or between self and entity, and because they are highly frequent in the input. When the children did start to use verbs they used primal verbs to a significantly greater extent than other verbs. When they began to combine words, their reliance on relational words alone to express location or motion decreased as primal verb use increased. Together, relational words and primal verbs, used alone or together, comprised the dynamic action component in an average of 85% of the children's total word combination types during the first several months of combinatorial language.

So how can this selection of dynamic event words as well as the relative timing of each type be explained in a developmental context? In the single word period, children use dynamic event words to express the basic predicative structure (an entity and its location or movement). These

meanings are also expressed in their sensorimotor actions with objects. One word expresses the entire motion situation but, given the child's representational ability, the only aspects that are being differentiated within the motion situation are the object versus its status in terms of location and time.

In the word combining period the child is able to mentally represent a more differentiated view of the motion situation and has an added interest, namely her own role (or the role of another) in changing the state or location of the object. The focus, at this point, is not only on how the object or the agent is affected by the motion. Rather the primary interest lies in the relationship between agent and object – the conveyance of the object to the agent (even **into** the agent in the case of consumption verbs) or the extension of an object from the agent to some point in space outside of the agent. Children's metaphorical ability is also evident in such primal verbs as *see*, connoting metaphoric ingestion, or *make*, connoting creation or metaphoric expulsion from the self (Ninio 1999a; Lakoff 1993).

The aspect of the motion situation being expressed here is deixis – direction with respect to the speaker. This emphasis is not surprising when one considers that the period of development in which children begin to produce early word combinations corresponds with important milestones in the development of self-knowledge (Bertenthal and Fischer 1978; Amsterdam 1972; Lewis et al. 1989). The self or other as agent is now prominent along with aspects of deixis, figure/ground relationships and path. At this point, the range of verb meaning expressed productively remains limited. Specific actions or movements seem not to be sufficiently differentiated from the agent, the object, and changes in their actual or metaphoric temporal and spatial relationships to permit learning of a wide range of more specific verbs.

In conclusion, we have shown that dynamic event meaning in English is first expressed with relational words, directly related to prior cognitive understanding and can be conceptualized as referring to motion event elements, that first verbs tend to be the set of primal verbs that overlap in meaning with relational words, and that the early combining period is dominated by relational words and primal verbs, collectively termed dynamic event words. These findings satisfy our goal of providing an initial demonstration that dynamic event word use provides a cognitive/semantic bootstrap to verb meaning.

ELLEN HERR-ISRAEL
Rutgers University, New Jersey

LORRAINE MCCUNE
Rutgers University, New Jersey

10. NOTES

[1] Single dynamic event words in English were first termed "function words" (Bloom, 1973); later, "relational words" (Bloom and Lahey, 1978). The term "relational word" is still useful in English because single dynamic event words tend to be particles in adult language, while in other languages may take the part of speech form of verbs. In contrast with entity words verbs are considered to be dynamic event words.

[2] Ninio's notion that children gain experience with a first, 'pathbreaking' group of verbs before a significant increase in production of new verb types corresponds with Ingram et al. (this volume) in their finding of a verb acquisition milestone when verbal syntactic types constitute 50% or more of the child syntactic types.

[3] For participants with double entries, E=English, S=Estonian, G=German.

[4] This word was invented by the children.

[5] This total includes all combinations, although routinized and imitated utterances were excluded from the analysis.

[6] Alice's utterances were analyzed to and including 20 months rather than 24 months because of her precocity relative to the other two children.

[7] For this analysis, Alice's utterances were analyzed to and including 20 months rather than 24 months because of her precocity relative to the other two children.

[8] Words in bold were also identified by Ninio.

[9] Words with an asterisk were used by all three children.

[10] Words in parentheses are non-primal verbs.

11. APPENDIX

Dynamic Action Words Used in the Single Word and Word Combining Periods[10]

Aurie

Single Word Period:

16 months: *down, open, up*
17 months: *off, stuck*
18 months: *no, out, (push)*

Word Combining Period

20 months: *get, (hit)*
21 months: *do, down, here, like, look, no, on, open, where, (bend, button, sleeping, tap)*
22 months: *allgone, eat, feel, help, hold, in, more, off, play, want, (talk, turn)*
23 months: *can, close, come, find, out, stuck, there, (tie)*
24 months: *can't, drink, found, go, have, leave, put, need, see, take, (calling, crying, hate, kiss, let, pat, read, spread, untie)*

Sandra

Single Word Period:

16 months: *back*

17 months: *eat, no, stuck*

18 months: *allgone, down, up* (*get, give, go* used in combinations)

19 months: *more, uhoh*

Word Combining Period:

20 months: *there, (sit, tickle)*

21 months: *another, come, do, got, here, hide, off, on, out, take, (bite, dump, fall, read, sleep, swing, turn, wipe)*

22 months: *away, feed, help, hold, need, open, other, over, put, want, where, (wrap)*

23 months: *close, down, eat, found, gave, like, make, move, no, play, putting, see, (brush, keep, sit, smell, sweep)*

24 months: *have, look, (call, comb, sweeping, swim, wear)*

Alice

Single Word Period:

16 months: *down, (clean)*

17 months: *empty, here, stuck, (wipe)*

Word Combining Period

18 months: *down, feel, here, no, see, (buy, drive, fix, ride, sit, talk, try)*

19 months: *another, close, come, coming, drink, eat, feel, find, fit, found, get, go, gonna, have, help, hold, look, make, more, play, off, on, open, out, stuck, take, there, up, want, where, (caught, clean, cut, hammer, lay, sleep, touch, wear)*

20 months: *at, can't, feed, leave, like, not, put, took, (catch, fill, iron, rock, say)*

12. REFERENCES

Amsterdam, Beulah. "Mirror Self-Image Reactions Before Age Two." D*evelopmental Psychobiology*. 5.4 (1972): 297-305.

Bertenthal, Bennett and Kurt Fischer. "Development Self-Recognition in the Infant." *Developmental Psychology* 14.1 (1978): 44-50.

Bloom, Lois. *One word at a time.* Mouton: The Hague, 1973.

Bloom, Lois and Margaret Lahey. *Language development and language disorders.* New York: John Wiley, 1978.

Bornstein, Mark, Cote, Linda R., Maital, Sharone, Painter, KKathleen, Pascual, Lilliana, Pecheux, Marie-Germain, Ruel, Josette, Venuti, Paola, and Vyt, Andrew Cross-linguistic analysis of vocabulary in young children: Spanish, Dutch, French, Hebrew, Italian, Korean and American English. *Child Development* (2004) 75: 1-25.

Bowerman, Melissa. "Mapping thematic roles onto syntactic functions: are children helped by innate linking rules?" *Linguistics* 28 (1990): 1253-1289.

Bowerman, Melissa. "Systemizing semantic knowledge: Changes over time in the child's organization of word meaning." *Child Development* 49 (1978): 977-987.

Bowerman, Melissa. "Learning how to structure space for language: A cross-linguistic perspective." In Paul Bloom et al. (eds.), *Language and Space*. Cambridge, MA: MIT Press, 1994: 385-436.

Bowerman, Melissa and Soonja Choi. "Shaping meanings for language: Universal and language-specific in the acquisition of spatial semantic categories." In Melissa Bowerman and Stephan Levinson (eds.), *Language Acquisition and Conceptual Development.* Cambridge, UK: Cambridge University Press, 2000: 475-511.

Braine, Martin. "Children's first word combinations." *Monographs of the Society for Research in Child Development*. Chicago: University of Chicago Press, 1976, 41.

Braine, Martin. "The ontogeny of English phrase structure: the first phase." *Language* 39 (1963): 1-13.

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

Choi, Soonja and Melissa Bowerman. "Learning to express motion events in English and Korean: the influence of language-specific lexical patterns." *Cognition* 41 (1991): 83-121.

Clark, Eve. "Strategies for communicating." *Child Development* 59 (1978): 953-959.

Gentner, Dedre. "On relational meaning: the acquisition of verb meaning." *Child Development* 49 (1978): 988-998.

Hopper, Paul and Sandra Thompson. "Transitivity in grammar and discourse." *Language* 56 (1980): 251-299.

Lakoff, George. "The contemporary theory of metaphor." In Andrew Ortony (ed.), *Metaphor and Thought.* 3d Edition. Cambridge, UK: Cambridge University Press, 1993: 202-251.

Leopold, Werner. "Speech development of a bilingual child: a linguist's record. Vol. I: Vocabulary Growth in the First Two Years." *Northwestern University Studies in the Humanities* 6 (1939): *xiv* – 188.

Lewis, Michael, Margaret Sullivan, Catherine Stanger, and Maya Weiss. "Self-development and self-conscious emotions." *Child Development* 60 (1989): 146-156.

Lieven, Elena, Julian Pine and Gillian Baldwin. "Lexically-based learning and early grammatical development." *Journal of Child Language* 24 (1997): 187-219.

Lyons, John. *Introduction to theoretical linguistics*. Cambridge, UK: Cambridge University Press, 1968.

McCune, Lorraine. "Dynamic event words: From common cognition to varied linguistic expression." First Language: (In Press).

McCune, Lorraine. "A normative study of representational play at the transition to language." *Developmental Psychology* 31 (1995): 198-206.

McCune, Lorraine and Marilyn Vihman. *Relational words + motion events: a universal bootstrap to syntax?* SRCD Poster. Albuquerque, NM, 1999.

McCune, Lorraine and Marilyn Vihman. "The transition to reference in infancy." Unpublished Ms., 1997.

McCune-Nicolich, Lorraine. "The cognitive basis of early relational words." *Journal of Child Language* 8 (1981): 15-36.

McCune-Nicolich, Lorraine and Carol Bruskin. "Combinatorial competency in symbolic play and language." In Kenneth Rubin (ed.), *The play of children: current theory and research*. Basel:Korger, (1981) 5-22.

Nicolich, Lorraine. "Beyond sensorimotor intelligence: assessment of symbolic maturity through analysis of pretend play." *Merrill-Palmer Quarterly* 23 (1977): 89-101.

Ninio, Anat. "Pathbreaking verbs in syntactic development and the question of prototypical transitivity." *Journal of Child Language* 26 (1999a): 619-653.

Ninio, Anat. "Model learning in syntactic development." *The International Journal of Bilingualism* 3 (1999b): 111-131.

Piaget, Jean. *The Construction of Reality in the Child.* New York: Basic Books, 1954.

Pine, Julian and Elena Lieven. "Reanalyzing rote-learned phrases: individual differences in the transition to multi-word speech." *Journal of Child Language* 20 (1993): 551-571.

Pinker, Steven. *Language learnability and language development.* Cambridge, MA: Harvard University Press, 1984.

Sinclair, Hermina, Mira Stambak, Irene Lezine, Sylvie Rayna, and Mina Verba. *Infants and objects: the creativity of cognitive development.* New York: Academic Press, 1989.

Talmy, Leonard. " Toward a cognitive semantics." *Concept Structuring Systems.* (Vol. 1). Cambridge, Mass.: MIT Press, 2000.

Talmy, Leonard. "Lexicalization patterns: semantic structure in lexical forms." In Timothy Shopen (ed.), *Language typology and syntactic description*. Vol. III: Grammatical categories and the lexicon. Cambridge, UK: Cambridge University Press, 1985: 36-149.

Talmy, Leonard. "The semantics and syntax of motion." In John Kimbal (ed.), *Semantics and Syntax.* New York: Academic Press, 1975: 181-238.

Tomasello, Michael. *First verbs*. New York: Cambridge University Press, 1992.

Veneziano, E. (2003). "The emergence of noun and verb categories in the acquisition of French". *Psychology of Language and Communication,* 7 (1), 23-36.

Vihman, Marilyn. "The transition to grammar in a bilingual child: positional patterns, model learning, and relational words." *The International Journal of Bilingualism* 3.2, 3.3 (1999): 267-301.

Vihman, Marilyn and Ruth Miller. "Words and babble at the threshold of language acquisition." In Michael Smith and John Locke (eds.). *The emergent lexicon: The child's development of a linguistic vocabulary.* Developmental psychology series. San Diego, CA: Academic Press, 1988: 151-183.

Werner, Heinz and Bernard Kaplan. *Symbol Formation*. New York: Wiley, 1963.

DAVID INGRAM, ANNE WELTI AND CHRISTINE PRIEM

THE EARLY STAGES OF VERB ACQUISITION IN SPANISH, GERMAN AND ENGLISH

Abstract. This article provides a preliminary report on how children acquire verb paradigms, proposing a set of verb stages. The stages are: 1. Holophrases: the first action words appear; 2. Early Word Combinations: actions words combine to form semantic relations; 3. Early Verbal Combinations: verbs become obligatory, marked by an increase in verb use, and a later increase in the number of verb types; 4. Early Paradigms acquisition of inflected forms of individual verb types. A method is presented that places individual children into these stages. The method uses 50 intelligible utterances and measures the frequency of verb types, verb forms, and verbal syntactic types, i.e. word combinations that contain a verb. The method is applied to data from Spanish children with language delay, one typically developing German child, and two typically developing children acquiring English. The analyses support the stages, and show there is a milestone of verb acquisition when verbal syntactic types increase to over 50% of the child's syntactic types. It is also shown that early verb acquisition is marked by the acquisition of verb types, followed by the gradual addition of verb forms (and thus paradigms) over an extended period of time.

1. INTRODUCTION

The importance of verbs in first language acquisition has been recognized from the earliest diary studies (e.g. Stern and Stern 1907) to the more systematic longitudinal and experimental studies of the 1970's (c.f. Bloom 1991), to the current era of child language studies (Tomasello 1992). It has been observed that verbs do not constitute the first class of words acquired, but rather emerge somewhat gradually after an early period of lexical growth focused on nouns (Stern 1924). Once verbs become more common, there is a subsequent extensive period of acquisition when verb inflections are gradually acquired (Brown 1973). It is also noted that the emergence of verb inflection is piece-meal during the first months, i.e. that specific verbs are restricted in the number of inflections that occur with them (Gathercole, Sebastian and Soto 2000; Ingram and Thompson 1996).

Despite these findings, there are at least two areas where our knowledge is still limited, and these two areas will be the focus of this article. First, we have

N. Gagarina and I. Gülzow (eds.), The Acquisition of Verbs and their Grammar: The Effect of Particular Languages, 151–171.

yet to formulate precise stages of verb acquisition. The primary research that has attempted to do this has speculated on when children recognize that verbs are required, particularly when tense is required. This line of work has proposed two stages of verb acquisition, a first optional infinitive stage, when children do not require verbs to be tensed, followed by a second stage when tensed verbs are required in main clauses (Wexler 1994). The details of this proposal have been presented in a series of articles by Wexler and his colleagues, and have been controversial (Ingram and Thompson 1996). It is not the purpose of the present study to critique this approach, however, but to provide an alternative, more general, model of verb acquisition.

The second area to be addressed concerns the acquisition of paradigms. Work on the acquisition of verb inflections from the seminal work of Brown (1973) has treated verb inflections as independent entities. For example, Brown provides findings on English verb inflections indicating an order of acquisition such that –ing tends to be the first one acquired and –ed noticeably later. Linguists for many years, however, have known that verb inflections are structured into paradigms, and that paradigms can be highly structured (Williams 1994; Wurzel 1987). The goal of language acquisition studies, then, should be an understanding how particular paradigm classes are acquired, not just the specific endings. The model of verb acquisition to be presented therefore incorporates the notion of verb paradigm.

The acquisition of verbs is central to the understanding two other topics in language acquisition, individual variation between children and how grammar is acquired in general. If verbs are key to grammatical acquisition, then one should anticipate their stages of acquisition to be relatively uniform across children. Also, if children use the semantics of action words to bootstrap into syntax (c.f. Pinker 1984), this claim gains support if empirical data on verb acquisition indicate a 'verb spurt' or discontinuity of some kind.

In addition to the issues of verb stages and paradigms, the article also will address methodological aspects of the study of verb acquisition. It will be demonstrated that children in the early stages of grammatical acquisition can show gains in verb usage without noticeable changes in MLU. It is important, therefore, to have a procedure for determining a child's verb stage that is not tied to MLU. A procedure for doing this will be presented, based on the importance of determining verbal syntactic types and controlling for sample size. The other methodological area to be discussed is how language samples are recorded, transcribed, and stored. A procedure is outlined for this as well, incorporating new personal computer technology for the digitalization of language samples.

The actual presentation is as follows. The first section provides a model for stages of verb acquisition, incorporating the notion of verb paradigms. The next sections give a procedure that operationalizes the proposed stages. One section discusses how the use of digitalized sound files can provide better data for the study of verb inflections. This is followed by a section on the procedure for placing a child into the specific verb stages. The last sections present preliminary analyses of data from children acquiring Spanish, German and English; the data demonstrate that the proposed stages are able to capture similarities across children cross-linguistically and in turn provide evidence for the viability of the stages proposed. The data also show some limited individual variation, and evidence for a shift to a stage in which verbs are central to the grammar.

2. STAGES OF VERB ACQUISITION

The following verb stages for the early acquisition of verbs are based on the broad literature on language acquisition since the earliest studies over 100 years ago (c.f. review in Ingram 1989). They are built from some basic assumptions about how grammatical acquisition proceeds from the first words to sentences, and are not founded on any particular theory of syntax. The stages are, however, founded on the well accepted linguistic notion that verbs form paradigms. The challenge then, is to formulate stages that follow the language learner from the acquisition of the first instances of individual verb forms to their eventual incorporation into paradigms.

In order to acquire verbal paradigms, children need to know each of the following:

1. *language consists of linguistic signs*, i.e. pairings of meanings with phonetic sequences. The early meanings cover words that refer to a wide range of concepts such as objects, people, actions, attributes, and personal-social actions (Nelson 1973). This milestone typifies much of the acquisition that takes place during the holophrastic period. Action words constitute a part, but not a large part, of what is acquired during this period. Apart from speculation about semantic bootstrapping (Pinker 1984), there is little evidence per se in the speech of children at this point that a category of verb exists (c.f. also Olguin and Tomasello 1993).

2. *language consists of word combinations*, i.e. pairings of words that themselves form a higher unit of organization. Studies in the 1970s found that the early word combinations of children were semantically consistent and based on a

basic set (Brown 1973). The exact nature of these combinations is controversial (i.e. semantic vs. syntactic), but a shared assumption of most if not all accounts is that they involve combinatory knowledge. A conservative view of this knowledge is that the hierarchical knowledge is relatively flat (Brown 1973; Hill 1984), such that the occurrence of 'agent + action' and 'action + object' does not necessarily imply a higher structure such that a sentence consists of a subject noun phrase and a predicative verb phrase;

3. *language consists of word combinations that require a verb*, i.e. that predication forms the basis of sentences. The understanding that the verb is the central unit of grammatical structure is an important milestone and one that can take place without necessarily having an increase in sentence length. There are two expected events that should take place when this milestone is reached. First, one would expect a noticeable increase in the number of word combinations containing a verb. Second, there should be an increase in the number of different verbs used by the child, as the child recognizes their importance. The occurrence of such changes would constitute evidence that the semantic bootstrapping of verbs has taken place.

4. *language consists of verb paradigms*, i.e. sets of verb forms that share a core meaning and vary in their expression of time, aspect, person and number. The point when children reach this milestone could be quite a bit later than the earlier developments, given the number of verbs that exist and the number of available verb forms. It is anticipated that children may acquire a relatively large number of verb types before adding a range of inflected variants to them (Gathercole, Sebastian and Soto, 2000). While the process of paradigm building would be expected to be similar across children, there is nothing about the process that precludes individual differences between children concerning the specific intermediate paradigms that are created.

The proposal is that children acquire the above linguistic principles in the order presented, resulting in four early stages of verb acquisition given in Table 1. There are four major stages proposed, holophrases, early word combinations, early verbal combinations, and early paradigms. The third stage of Early Verbal Combinations consists of two substages, one when the rate of verbal word combinations increases, followed by an increase in the use of different verb types. It is expected that future research will adjust these stages, particularly the last stage of verb paradigms, which likely consists of substages as well.

Table 1. Hypothetical stages in the acquisition of verb paradigms

1. Holophrases: children acquire their first words that cover a wide range of meanings; one category of meaning are action words that will later bootstrap into the category of verb; at this stage action words constitute a small number of the child's total vocabulary;
2. Early Word Combinations: children recognize that words can combine into categories that cover a range of possible semantic relations, e.g. possessor possessed, action object; action words at this stage are just one of the possible combining units and do not occur at a higher rate than other possible words;
3. Early Verbal Combinations: children recognize that verbs (or action words) constitute the central unit for word combinations; there is an increase of the number of word combinations containing verbs (3a), followed by a concomitant increase in the range of verbs used (3b); these increases mark the point at which semantic bootstrapping has occurred;
4. Early Paradigms: verb acquisition proceeds to a point where the child acquires both a vocabulary of different verbs but also variant inflected forms of verbs as well; this stage of development will cover an extended period of time, and may itself consist of substages, with individual variation across children.

3. LANGUAGE SAMPLING

In order to test the viability of the proposed verb stages, it is necessary to resolve methodological issues around language sampling and the determination of operational definitions of the stages. Data to examine the stages can come from both databases of language samples such as Childes (MacWhinney and Snow 1985), and the collection of new samples. The data provided later is this article uses both, with the majority coming from the latter source.

Despite the availability of growing databases on child language, there is a fundamental concern in using such data for the study of verb acquisition. The fundamental reason is that verb inflections across languages tend to be particularly difficult to transcribe. They are often nonsyllabic (particularly in English), at the ends of words, and not necessarily very frequent. For example, the sample to be discussed later from Eve (Brown 1973), found only eight instances of past tense in a sample of 1000 utterances. The introduction of transcriptional error in such cases could influence greatly claims about the acquisition of specific grammatical inflections.

Given these concerns, the approach we advocate is a combination of such databases with data collected first hand, digitized, and transcribed using recent advances in personal computer technology. These advances enable us to provide transcriptions that will likely be more accurate and are accessible to independent verification.

The method of sampling that we follow is outlined in more detail in Bunta, Ingram and Ingram (2003). Language samples from children are recorded either directly into digital format or converted into digital format. The resultant digital sound files are then edited through computer manipulation in a way that each child utterance and surrounding adult utterances are extracted as its' own sound file. This process enables transcribers to listen to individual utterances directly on the computer where they can see the actual wave form of each utterance, play them slower and or louder as needed, and easily repeat them. It also provides easy access to any utterance at any point, since the individual sound files can be put onto CDs and shared and transported easily. This enables the researcher to do independent transcription on all utterances, and verify with spectrographic analysis if necessary to determine the presence of grammatical morphemes and their acoustic properties.

4. VERB STAGE ANALYSIS

Regardless of the procedure followed to obtain a language sample, it is also necessary to have a way to test the proposed stages by placing the child into one of the four stages in Table 1. Developing such a procedure involves decisions about sample sizes, defining verb and verb forms, and measuring the occurrence of verbs and their paradigms. This section discusses a procedure that has proven capable thus far to be capable of doing this.

4.1. Language Sample

Typically, a language sample should be one that is large enough that it captures what is being sought without being so time consuming as to be unpractical. Given the potential infrequency of some verb forms and endings, it might initially be anticipated that language samples would be need to be quite large. At the same time, it is possible to get an idea of a child's verb usage from shorter samples if sample size is held constant and measures are made in relation to that sample. For example, if the majority of a child's sentences contain a verb, this can be determined from a smaller sample even though the full range of verbs would not

be. This approach also enables larger samples to be included, and long as they are divided into the same sizes, with measures averaged across the samples.

In addition to sample size, there is the related issue of deciding what utterances to include. Since unintelligible utterances contain little useful information, these are commonly deleted from analyses. Other possible candidates for deletion are self-repetitions (the child saying the same thing twice) and imitations (of the preceding adult utterance). The flip side of excluding utterances is that determining whether an utterance meets such an exclusionary category can be time consuming. For example, defining an imitation can be a difficult process, particularly in terms of how distant the child's imitation needs to be. An efficient procedure will be one that allows for quick and reliable decisions on what is or is not included.

Another issue about language sampling is decisions about having a sample that reflects linguistic performance or linguistic competence. For example, counting all utterances will reflect linguistic performance, while excluding a large number of utterances leans in the other direction. Developmental Sentence Scoring (Lee 1974), for example, selects 50 unique utterances containing a subject and a verb. For children at earlier stages of acquisition, a substantially large sample would be needed to obtain 50 unique utterances.

The present procedure selects 50 intelligible utterances with some exclusionary criteria that are intended to be reliable and to provide a valid sample. Since the focus is on grammatical acquisition, utterances are only included if they contain a word or words that either enter or potentially can enter into a grammatical relation. A holophrastic utterance like "eat" would be included, but one like "yes" would not. Besides excluding fillers like "oh", etc. other single word utterances are excluded if they contain a proper noun such as "mommy", since such occurrences could be vocative usages and not part of a grammatical structure. Excluding these words eliminates the potential of having a sample consisting of many instances of a child calling their mother. Immediate self-repetitions are excluded, i.e. the child repeating himself exactly, but non-immediate ones are not. Imitations are also not excluded. The rationale is that the latter two types of utterances can be time consuming to identify.

4.2. Verbal categories

Once a 50 utterance sample is determined, the following verbal categories are determined and counted. The number of lexical verbs is determined for the sample, divided into *verb types* i.e. distinct verbs in meaning ("eat" vs. "drink"),

and *verb forms* (all inflected or non inflected forms). For example, "eat" and "eating" are one verb type but two verb forms. Next, the number of syntactic types is counted (Bloom, Lightbown and Hood 1975). A *syntactic type* is a unique utterance that contains at least two words that enter into a grammatical relation. For example, "eat cookie" with an action object relation is a syntactic type, but "eat, mommy" with "mommy" used as a vocative is not. Syntactic types are then subdivided into those that contain a verb, i.e *verbal syntactic types*, and those that do not contain a verb, i.e. nonverbal syntactic types. Lastly, a verb paradigm is determined. A *verb paradigm* is a grouping of verb types that have verb forms showing the same verb inflections. Table 2 summarizes the procedure followed. The following sections demonstrate the application of the procedure to language samples from Spanish, German and English to test the proposed stages.

Table 2. Procedure for determining verb stages

- 50 successive intelligible utterances selected for further study;
- exclude single word utterances that contain a proper noun, e.g. "mommy", "daddy", since they may be functioning as vocatives;
- exclude fillers such as "oh", "uhhuh" since these do not enter grammatical relations;
- exclude common forms like "yes", "no" since they may be frequent and not reflect grammatical usage;
- exclude immediate self-repetitions;
- **intelligible utterance** is one that contains at least one word that enters into a grammatical relation; eliminates words like "yes", "okay"; "uhhuh", etc.
- determine number of **lexical verbs** (types and forms), e.g. "hold" " holds" is one verb type (the verb "to hold") and two verb forms;
- determine the number of syntactic types, i.e. a unique utterance with at least two words that enter into a grammatical relation, e.g. "eat cookie" vs. "eat, mommy"
- determine the number of **verbal syntactic types (VST)** i.e. syntactic types that contain a lexical verb;
- determine **verb paradigms**, i.e. groupings of verbs that share the same verbal affixes.

5. EARLY VERB STAGES IN SPANISH

5.1. Data

Ingram, Kayser, and Durfee (2002) analyzed the early verb stages of five Spanish speaking children attending a bilingual preschool in the American Southwest, and identified as having a language delay. Language delay was defined as having an MLU of approximately 2.0 and being between 42 and 54 months of age. These MLU values are lower than any of the children identified as language delayed in Bedore and Leonard (2001) who were also in this age range, and in some cases younger. Spontaneous language samples were recorded of approximately 30 minutes in length and analyzed as explained above. All child utterances and surrounding adult utterances were extracted as sound files. Independent transcriptions were done on all utterances. Table 3 provides the general sample information for the children studied, along with the results of the analysis.

The first finding to note is that all the children produced syntactic types, ranging from 27 to 38, indicating that they were at least in a stage that involved word combinations. The children differed, however, in their use of verbal syntactic types. The two children with the lowest MLUs, R and D, also had the lowest rates of verbal syntactic types at 30% and 19% respectively of the total number of syntactic types. The other three children had percentages of verbal syntactic types at over 50%. This distinction of below or above 50% verbal syntactic types provides a potential measure to distinguish between children the Early Word Combination stage and those at the Early Verbal Combination stage.

Table 3. Participant information for five children with language delay reported in Ingram, Kayser and Durfee (2002)

	Children				
Measures	R (female)	D (male)	L (male)	A (male)	J(male)
Age	3,8;27	3,9;1	4,9;17	4,5;24	4,7;27
Sample Size	170	233	316	202	100
MLU (words)	1.69	1.72	1.90	1.96	2.17
Intelligibility	41%	36%	61%	57%	56%
Verb Measures					
Synt. Types	30	27	28	38	28
% VSTs	30%	19%	54%	55%	61%
Verb Types	8	6	9	16	15
Verb Forms	9	7	10	19	21
Stages	2	2	3a	3b	3b

Turning to the distribution of verb types, it is noteworthy that the number of verb types for three of the children, R, D, and L, was much lower than the other two children A and J. Child L is of particular interest because he showed a rate of verbal syntactic types similar to the more advanced children, but a rate of verb types similar to the two least advanced. His results are suggestive of stage 3a, i.e. a substage of the Early Verb Combination stage where there is an increase in the use of verbal syntactic types without an increase in the number of verb types. The distribution of the verb types suggests a second heuristic for the verb stages, with children having fewer than ten verb types being in the first stage of Early Verbal Combinations (3a) and children with more than ten verb types falling into later stages.

The last measure applied to the data was the number of verb forms. For three of the children, R, D, and L, there was only one more verb form than verb types, indicating that most verbs followed the pattern of one form per verb type. The child with the most forms, J, did have 6 more forms that verb types. A closer examination of his verbs, however, indicated that 11 of the 15 verbs only had one form, and several of these were irregular forms, e.g. "saber" (know) /sé/, "ser" (be) /es/, "tener" (have) /tengo/. Of the remaining 4 verbs with multiple forms, "estar" (be) had four of them, i.e. /estaba/, estabas/, /estan/, /está/. Given these numbers, it is evident that there is still very little productivity in J's verb system.

In examining verb forms, the question arises concerning how to define productivity in relation to the existence of a verb paradigm. One suggestion on how to do this is found in Gathercole, Sebastian and Soto (2000). They propose a very minimal requirement that an inflection is considered productive if it is produced in at least two verbs and one of those verbs is found with another inflection. Applying this criterion to J's verbs yields the patterns in (1):

(1) "brincar" (to buck)	"mirar" (look at)	
/brincó/		3rd past perfect
/brinca/	/mira/	3rd p. sg. Present

This yields the verb morpheme /-a/, and a very small emerging paradigm class of two verbs. Even by this weak criterion there is little evidence of paradigmatic structure for J, the most advanced of the five children studied. The analysis supports previous studies is that children acquire their first verbs in a piece meal fashion with little evidence of paradigmatic structure.

The data from Ingram, Kayser and Durfee (2002) provide some preliminary measures for placing the children into the stages proposed in Table 1. The suggestive measures are summarized below (where VST = verbal syntactic type), and they will be applied and assessed in the next sections:

(2) *Possible Measures for Early Verb Stages (based on 50 utterance samples)*

1. Holophrases: single word utterances with a low rate of verb use (under 10); more specific measures to be determined;
2. Early Word Combinations: VSTs less than 50% of syntactic types, verb types under 10;

3a. Early Verbal Combinations I: VSTs increase to over 50%, without an increase in verb types;

3b. Early Verbal Combinations II: increase in VSTs followed by an increase in verb types to between 10 and 20, with no comparable increase in verb forms;

4. Early Paradigms: number of verb forms increase to a point when paradigms begin to emerge; specific measures to be determined.

5.2. Summary

The data from the Spanish children with language delay indicated that the method of using 50 intelligible utterances and measuring verb types, forms, and verbal syntactic types provided the potential for determining a child's level of verb acquisition. The analyses indicate the following general conclusions about verb acquisition: 1. There is a milestone of verb acquisition occurs when verbal syntactic types increase and constitute over 50% of the child's syntactic types; 2. Early verb acquisition was marked by the acquisition of verb types, later followed by the gradual addition of verb forms, and 3. Verb paradigms are acquired over a long period of time, and at some point to be determined. These findings will be explored further in the next sections with data from case studies on German and English.

6. EARLY VERB STAGES IN GERMAN

6.1. Data

Ingram and Thompson (1996) studied the acquisition of early verb inflection in four German children between 1;5 and 2;5, three from the CHILDES data base and

one child, Dorothy, from their own. Their general conclusion was that a strict criterion revealed little productivity in the children's verb systems, with only two tense markers showing any kind of productivity, the first person singular /-0 or -e/ and the third person singular /-t/. They also noted "Only four verbs are marked for more than one tense, and only two of these, "kommen" and "gehen" appear in the samples of more than one child" (p. 111).

The present report returns to the database from Dorothy to examine her verb stages and verb paradigms from the methodology presented above. Dorothy is a German child followed longitudinally in 13 language samples, approximately one hour each, two to three weeks apart, between 1,11 and 2,7. While German was the primary language in the home, she lived in an English speaking community (Vancouver, Canada), and as will be shown, English began to appear in her speech toward the latter sessions. (As mentioned, a subset of these data are reported in Ingram and Thompson 1996). Language samples were digitized and individual child utterances extracted for subsequent transcription and analysis. 50 utterance samples were determined and analyzed for verb use as described for each sample. Samples were then grouped in threes (and in one instance four) to compress the data. Analyses were then conducted measuring the distribution of verbs and verbal syntactic types as described above for the Spanish children. The main difference from the Spanish study is that the data for Dorothy were quite larger, with 650 total utterances examined.

Table 4 presents the main information about Dorothy's samples and the results of the analyses. A first observation to make is that Dorothy's 4 time periods can be subdivided into two periods, the first two being at the end of the holophrastic period (MLUs of 1.24 and 1.36), and the latter two periods showing a higher rate of word combinations (MLUs of 1.85 and 1.82). This distinction is based on the recommendation made in Ingram (1989), that a child near the end of the holophrastic stage produce sequences of single word utterances that can be misidentified as word combinations. A child is not credited as being in the stage of word combinations until there production reaches a higher rate of frequency as either measured by the overall number of word combinations or by MLU (where MLUs over 1.5 suggest that word combinations are beginning to out number holophrases). The overall rate of producing syntactic types does not increase dramatically for Dorothy, but the increase in MLU in the latter sessions supports the notion that she has reached a stage of word combinations.

Table 4. Dorothy's acquisition of German verbs (based on 20 utterance samples)

	Time Periods (no. of samples in parentheses)				
Measures	*1 (3)*	*2 (3)*	*3 (3)*	*4 (4)*	*Total*
Age	1,11;27	2,1;24	2,5;7	2,6;22	
MLU (words)	1.24	1.36	1.85	1.82	
Number of utterances	150	150	150	200	650
Syntactic Types	3	10	17	17	
VSTs	(67%)	50%	53%	47%	
Verb Types	1	5	6	8	
Verb Forms	1	6	7	9	
German					
New Verb Types	2	10	3	7	22
New Verb Forms	2	12	9	8	31
English					
New Verb Types	1	1	1	7	10
New Verb Forms	1	1	2	8	12

Dorothy's use of verbal syntactic types hovered around 50% throughout the period of observation, and did not show a marked increase at any point. Her rate of using verb types and verb forms was also relatively low, at or below ten throughout. Her data, therefore, place her in a period of word combinations, but do not break down into the finer distinctions of the stages proposed. She did not show a marked increase at some point in either the use of verbal syntactic types or verb types.

Some developmental changes can be observed concerning the emergence of new verb forms in relation to new verb types. The bottom of Table 4 gives these figures separated for Dorothy's German and English words. At time period 3, there was a marked increase in the appearance of new German forms, while the number of new verb types remained low. This is a pattern not seen in the Spanish data, but one to be expected when children begin to acquire paradigms. This increase in new forms, however, was not mirrored in the English verbs. The German verbs, however, suggest that Dorothy is beginning to acquire some primitive paradigms.

The analysis of Dorothy's verbs showed that most verb types still only had a single verb form. These verbs fell into two groupings, an infinitive class and a

finite class. These are shown in (3a) and (3b). Another group of five verbs had two forms, but there was little in the way of shared patterns (3c). Lastly there were three verbs that had three or more forms, as shown in (3d).

(3) *Dorothy's Emerging Verb Classes*
a. Infinitive class: "geben", "holen" , "laufen", "tun", "kennen", "aussleigen", "auszieben"
b. Finite class: "spielt", "weint", "weiss", "kann"
c. Two member class:
i.e. "essen" / "gegessen"; "geht" / "gegangen"; "habe" / "haben"; "schreibt" / "schreiben"; "ist" / "sind"
d. Verbs with three or more forms:
i.e. "muess", "musstest", "musste"; "komme", "kommt", "kommen"; "mache", "machst", "macht", "machen"

Dorothy showed more verb forms relative to verb types than the children in the Spanish data. At the same time, it is not obvious that her paradigmatic system is at anything but the most primitive level. Applying a minimal criterion of productivity such as used by Gathercole, Sebastian and Soto provides evidence for the following inflectional endings: infinitive /-en/, first person singular present /-0 or -e/, and third person singular present /-t/. If we apply a stricter criteria, however, that these endings need to appear at least twice on the same verb, then the evidence is limited to four verbs, i.e. "haben", "schreiben", "kommen" and "machen", as shown in (4):

(4) "haben"	"schreiben"	"kommen"	"machen"	
/habe/		/komme/	/mache/	1 p. sg. pres.
	/schreibt/	/kommt/	/macht/	3 p. sg. pres.
/haben/	/schreiben/	/kommen/	/machen/	infinitive

If this is accepted that an emerging paradigm, the data from Dorothy indicate that this stage can be reached without going through all the word combination stages of Table 1. A more conservative position could be that these are still not very productive forms and that stronger evidence is needed to say this is the case. The data also need to be cautiously interpreted because of several factors. For one, we are dealing with a single case study. Dorothy was not a particularly intelligible child, and it may have been that more word combinations were not transcribed as intelligible because they were harder to understand than single

word utterances. Another issue involves the fact that it is a different language, although one might expect that Spanish children might show earlier paradigms than German ones due to the richer morphological system. This points leads to the third factor that the Spanish children were ones with language delay, and such a delay may involve delays in paradigmatic acquisition. Lastly, this analysis involved more extensive data than the Spanish study.

6.2. Summary

The analysis of data from a German-speaking child only partially supports the stages proposed. They show that paradigms begin to be built up slowly after the acquisition of most verb types with a single verb form. They do not support the specific stages of word combinations in that they do not show stages of marked increases in verbal syntactic types or verb types, though they do show a stage where an increase in verb forms took place. The data therefore do not show a clear point at which semantic bootstrapping has taken place, and suggest that children may differ to the extent in which they overtly show such a transition.

7. EARLY VERB STAGES IN ENGLISH

The discussion in this section turns to two case studies in English. The first one is the first author's son Daniel (Ingram 1981), and the other child is Eve (Brown 1973), taken from the Childes database.

7.1. Daniel

Daniel was audio recorded approximately every two weeks from 1;6 to 2;11. Language samples were digitized and individual child utterances extracted for subsequent transcription and analysis. 50 utterance samples were determined and analyzed for verb use as described above for each of 18 samples. Samples were then grouped in threes to compress the data. This resulted in six sample periods, each with a sample of 150 utterances, for a total sample of 900 utterances. Measurements were averaged across the three samples at each sample period.

Table 5 presents the results of the analysis. In terms of MLU, Daniel's samples began at the holophrastic stage with an MLU of 1.04 and eventually crossed each of Brown's stages (Brown 1973) up to Stage II. His data provide a picture of development beyond that seen in either the Spanish or German data. The first two sample periods fell into the holophrastic stage (defined as an MLU

below 1.5 and a small number of verb types, under 5 per 50 utterance sample). The three samples at sample period 3 showed the characteristics of the Early Word Combination stage, i.e. word combinations, verb types below 10, and a low usage of verbal syntactic types (36%). Both these rates increased greatly for sample period 4, (14 and 66%), showing the transition to the Early Verbal Combination stage. Lastly, there was an increase of verb forms in periods 5 and 6 suggestive of the emergence of early paradigms. The data from Daniel, therefore, support the stages in Table 1.

Table 5. General information on Daniel, over 6 time periods

			Verbs		*Verbal Syntactic Types*	
Periods	*Age*	*MLU*	*Types*	*Forms*	*No.*	*%*
Holophrastic						
1. Presyntactic	1,7	1.04	1	1	1	33%
2. Early Stage I	1,10	1.39	4	4	5	48%
Early Word Combinations						
3. Late Stage I	2,1	1.67	7	7	7	36%
Verbal Word Combinations						
4. Early Stage II	2,4	2.03	14	15	17	66%
Early Paradigms						
5. Early Stage II	2,7	2.00	12	15	16	58%
6. Late Stage II	2,10	2.39	16	18	24	70%

Table 6 gives the emerging verb paradigms identified in Daniel's verbs. Most of Daniel's verbs (70%) occurred either in a bare form or with the /-ing/ progressive ending. As found in the previous analyses, most verb types occurred in just one form. The most common groups were bare verbs (26) or verbs with the progressive ending (13). There was a third group of verb types that only had a verb form in the past (one irregular and two regulars). The main paradigm found was a group of 18 verbs that had bare verb forms and ones with the progressive /-ing/. Two other classes were emerging: 1. an auxiliary class (8), and a small group of irregular verbs that contained at least one bare verb and an irregular past tense.

Table 6. Daniel's Emerging Verb Paradigms

Verb Types with One Verb Form:
Zero class (verbs with no verbal endings, 26 members), e.g. got, know, buy, draw, give, make, want, like, look, push, see, find, open, etc.
The Progressive class (verbs only found with /-ing/, 13 members), e.g. coming, dancing, drinking, falling, running, helping, sleeping, etc.
Past tense class: (verbs only occurring in past tense, 3 members), broke, squeezed, started
Verbs Types with Two Verb Forms:
The Progressive paradigm class (verbs that have both zero forms and ones with /- ing, 18 members), e.g. eat/eating, play/playing, put/putting, watch/watching, etc.
Verbs Types with Two or Three Verb Forms:
Irregular class (irregular verbs, 3 members), e.g. go/going/went, have/having/had; tell told
Other: Auxiliary class (8 members), e.g. can't, wanna, can, gonna, will

In summary, the data from Daniel supported the proposed stages of early verb acquisition. They also demonstrated the long process that verb paradigms follow in being acquired. Daniel's primary verb paradigm class was one that marked verbs as either bare verbs or verbs in the progressive. There was some evidence of separate classes emerging for auxiliaries and irregular verbs. There was little evidence of a productive past tense morpheme and no evidence of the regular third person morpheme.

7.2. Eve

The last case study to be examined is Eve from Brown (1973), whose data are available on the CHILDES data exchange (MacWhinney and Snow 1985). Eve's development was marked by very rapid expansion of her MLU at the expense of her acquisition of grammatical morphology, which lagged behind. She was also by far the fastest learner of the three children that Brown studied. An advantage of studying Eve is that her data provide information beyond Stage II, unlike the

above studies. 50 utterance samples were determined and analyzed for verb use as described above, for eighteen language samples from age 1;7 to 2,3. The MLUs ranged from 1.55 to 4.15. This resulted in a total of 900 utterances. As for Daniel, the samples were grouped in sequences of threes to create six sample periods of 150 utterances each.

Table 7 presents Eve's data in a way that can be compared to Daniel's data in Table 4. Eve was already in the stage of Early Verbal Combinations in the first two sample periods, which are comparable to Daniel's fourth sample period. The increase in verb forms in comparison to verb types in the following four periods leads us to identify these as a period of Early Paradigms. The other striking feature of these latter periods was an increase of verbal syntactic types to around 74%. Eve is clearly the most advanced of the children studied in the present article.

Table 7. General information on Eve, over 6 time periods

			Verbs		*Verbal*	
Periods	*Age*	*MLU*	*Types*	*Forms*	*Syntactic No.*	*Types %*
Verbal Word Combinations						
1.	1,7-1,8	2.62	10	10	21	58%
2.	1,8-1,9	2.73	13	14	31	60%
Early Paradigms						
3.	1,10-1,11	3.06	15	17	34	62%
4.	1,11-2,0	3.78	16	18	37	74%
5.	2,1-2,2	3.73	17	21	37	72%
6.	2,2-2,3	3.74	14	18	37	74%

The analysis of Eve's verbs combined all the samples, and included two samples excluded from the earlier analysis, for a total of 1000 utterances. It was found that there were overall 79 verb types and 113 verb forms, supporting the conclusion that there were emerging paradigms. Despite Eve's more advanced development, it was found again that most of her verb types (88%) only had one or two forms. The actual distributions are shown in (5):

(5) Percentage of Verb Types With:
1 form (65%) 2 forms (23%) 3 forms (10%) 4 forms (3%)

The distribution of verb forms by their inflectional endings is given in (6).

(6) Number (and Percentage of) Verb Forms:
bare (56%) -ing (25%) -ed (7%) past irreg (8%) 3rd reg (2%) 3rd irreg (2%)

Again it is seen that the variety of verb endings is very limited, with 81% of the verbs being either a bare verb or one with /-ing/.

The analysis of Eve's verbs into emerging paradigms resulted in a wider range of classes than found for Daniel's more limited verb system. These are presented in Table 8. As with Daniel, the main emerging paradigm was a progressive paradigm class, i.e. verbs that were either bare or took the progressive suffix. The emergence of past tense was more advanced, with both regular and irregular past tenses found with verb types having two and three verb forms, though these classes were still rather small. Eve's auxiliary class, however, was less developed than Daniel's. The fact that Eve's paradigms were still further underdeveloped supports the earlier claim that paradigm acquisition takes place over a very long period of language acquisition. Even though Eve had reached Brown's Stage V, she had not yet established English verb paradigms, and had an undeveloped auxiliary class. A comparison of Eve's and Daniel's data support the prediction that the specific intermediate paradigms that children create may differ, despite general patterns in the nature of the process of how paradigms emerge.

Table 8. Eve's Emerging Verb Paradigms

Verb Types with One Verb Form:
- Zero class (verbs with no verbal endings, 40 members)
The Progressive class (verbs only found with /-ing/, 9 members)
- Past tense class: (verbs only occurring in past tense, 2 members), forgot, ironed
Verbs Types with Two Verb Forms:
The Progressive paradigm class (zero forms and /-ing/, 10 members)
Past tense paradigm class (verbs that have both zero forms and ones with {-ed}, 4 members) fix/fixed; fold/folded; spill; spilled; see/seed
Irregular past tense paradigm class (verbs with zero forms and with irregular past, 3 members) break/broke; find/found; will/would

Verbs Types with Three Verb Forms:
Irregular class (irregular verbs, 3 members), e,g, make/making/made; write/writing/wrote; get/getting/got
Regular class (with {-ing} and past {-ed} 2 members): tie/tying/tied; cry/crying/cried

Verbs with Four Verb Forms: 3rd person class (one member): go/goes/going/gone

Other: Auxiliary class (2 members), e.g. can/can't; will/would

8. CONCLUDING REMARKS

This article has attempted to provide a preliminary report on one way to investigate how children acquire verb paradigms. It began by proposing a set of verb stages, based on known characteristics of children's language learning. A method of using 50 intelligible utterances and measuring verb types, forms, and verbal syntactic types was subsequently developed to test both the proposed stages and the method itself. The method proved insightful in characterizing the stages of Spanish children with language delay and two children acquiring English. It was less helpful in capturing the development of a German child, though it yielded information on her particular rate of acquisition for verbs and verb inflections. Her data suggest that the model may be capture some, but not all, children's pattern of verb development. Her data also suggest that not all children will overtly show evidence that semantic bootstrapping has occurred.

The analyses collective provided support that, for at least some children, there is a milestone of verb acquisition that occurs when verbal syntactic types increase and constitute over 50% of the child's syntactic types. The analyses also show that early verb acquisition is marked by the acquisition of verb types, later followed by the gradual addition of verb forms. Verb paradigms are acquired over a long period of time, and at some point to be determined after MLUs of 4.0. Despite this general pattern, the data from Daniel and Eve showed that the specific emerging paradigms that they develop will vary between children.

DAVID INGRAM
Department of Speech and Hearing Science
Arizona State University

9. REFERENCES

Bedore, Lisa., and Laurence Leonard. "Grammatical Morphology Deficits In Spanish Speaking Children With Specific Language Impairment." *Journal of Speech, Language, and Hearing Research* 44 (2001): 905-924.

Bloom, Lois. *Language Development: From Two To Three*. Cambridge: Cambridge University Press, 1991.

Bloom, Lois, Patricia Lightbown, and Lois Hood. . "Structure And Variation In Child Language." *Monographs of the Society of Research in Child Development* 40. 160: 1975.

Brown, Roger. *A First Language: The Early Stages*. Cambridge, Mass.: Harvard University Press, 1973.

Bunta, Ferenc, Kelly Ingram, and David Ingram. "Bridging The Digital Divide: Aspects Of Computerized Data Collection And Analysis For Language Professionals." *Clinical Linguistics and Phonetics* 17 (2003): 217-240.

Gathercole, Virginia, E. Sebastian, and P. Soto. "Lexically Specified Patterns In Early Verbal Morphology In Spanish." In M.R. Perkins, and S.J. Howard (Eds.), *New Directions in Language Development and Disorders*. (149-168). New York and London: Plenum, 2000.

Hill, Jane. (1984), "Combining Two-Term Relations: Evidence In Support Of Flat Structure." *Journal of Child Language* 11 (1984): 673-8.

Ingram, David. Early patterns of grammatical development. In R. Stark (Ed.) *Language behavior in infancy and early childhood*. New York: Elsevier North-Holland, 1981: 327-52

Ingram, David. *First Language Acquisition*. Cambridge: Cambridge University Press, 1989.

Ingram, David and William Thompson. "Early Syntactic Acquisition In Geman: Evidence For The Modal Hypothesis". *Language* 72 (1996): 97-120.

Lee, Laura. *Developmental Sentence Analysis*. Evanston, Ill.: Northwestern University Press, 1974.

MacWhinney, Brian and Catherine Snow. "The Child Language Data Exchange System". *Journal of Child Language* 12 (1985): 271-96.

Nelson, Katherine. "Structure And Strategy In Learning To Talk". *Monographs of the Society of Research in Child Development* (1973): 38-49 1973.

Olguin, R. and Michael Tomasello. "Twenty-Five-Month-Old Children Do Not Have A Grammatical Category Of Verb". *Cognitive Development*, 8 (1993): 245-272.

Pinker, Stephen. L*anguage Learnability and Language Development*. Cambridge, Mass: Harvard University Press, 1984.

Stern, William. *Psychology of Early Childhood Up to the Sixth Year of Age*. New York: Holt, 1924.

Stern, Clara and William Stern. *Die Kindersprache*. Leipzig: Barth, 1907.

Tomasello, Michael. *First Verbs: A Case Study of Early Grammatical Development*. Cambridge: Cambridge University Press, 1992.

Wexler, Kenneth. "Optional Infinitives, Head Movement And The Economy Of Derivations". In D. Lightfoot and N. Hornstein (Eds.), *Verb movement*. Cambridge: Cambridge University Press, 1994, 305-350.

Wurzel, Wolfgang. *Inflectional Morphology and Naturalness*. Dordrecht: Kluwer, 1987.

PETER JORDENS AND CHRISTINE DIMROTH

FINITENESS IN CHILDREN AND ADULTS LEARNING DUTCH

Abstract. This paper deals with the acquisition of finiteness in children acquiring Dutch as their first language and adults acquiring Dutch as their second language. The authors distinguish between the semantic concept of finiteness and its morpho-syntactic marking. Given that utterances are used to express illocutionary force, they argue that finiteness is the carrier of the pragmatic function of assertion. As such it relates the descriptive content of an utterance to its topic component. It is shown that for the expression of finiteness child and adult learners rely on this pragmatic function of assertion at subsequent stages of acquisition. At the so-called Conceptual Ordering Stage, i.e. before target-adequate morphological markings become productive, learners establish the assertive relation by a closed class of linking elements which contains elements expressing positive or negative assertion, modal phrases and scope particles. At the Finite Linking Stage assertion marking grammaticalizes. Elements of the target functional category of auxiliaries come to be used as a grammatical linking device whereas scope particles and other target adverbial elements do no longer occur as independent linking elements. While the illocutionary linking elements of the Conceptual Ordering Stage are adjuncts, auxiliary verbs are part of a functional category system. As is the case in the target language, they function as the head of a head-complement structure at the Finite Linking Stage.

1. INTRODUCTION

Finiteness is traditionally associated with morpho-syntactic properties of agreement and tense-aspect marking. In this paper however, we will investigate the acquisition of finiteness in children and adults learning Dutch from the point of view of the *semantic concept* of finiteness as defined in Klein (1998). It is claimed that the semantic properties of finiteness are expressed first. Furthermore, it is argued that auxiliary verbs serve as a bootstrap in the transition to the morpho-syntactic properties of finiteness. According to Klein semantic finiteness comprises two separate pragmatic or information-structural functions which can be termed as "anchoring" and "illocution". The anchoring function of finiteness provides the spatio-temporal co-ordinates into which the rest of the utterance is embedded; the illocutive function of finiteness validates the state of affairs expressed in the utterance with respect to these spatio-temporal co-ordinates.

N. Gagarina and I. Gülzow (eds.), The Acquisition of Verbs and their Grammar: The Effect of Particular Languages, 173–198.

In languages such as Dutch, German or English, both the anchoring and illocutive function are typically marked by one form, for example a finite copula as in the following example from Klein (1998:226).

(1) The book was on the table.

This fusion of two types of information in one form can be made explicit when contrastive stress is put on the copula as in (2).

(2) The book WAS on the table.

Unlike full verbs, copulae and auxiliaries do not carry any lexical meaning proper. Contrastive stress therefore only results in highlighting the two semantic components of finiteness. It can provoke both a tense contrast, and a contrast regarding the claim. There is a *tense contrast,* when (2) is opposed to *The book IS on the table*. What is highlighted here is the anchoring information, namely that the book's being on the table occurred at some moment in the past. There is a *contrast regarding the claim,* when (2) is opposed to *The book was NOT on the table*. What is stressed is the illocutive function, which expresses that the book's being on the table was indeed true for some spatio-temporal anchorpoint.

Both the anchoring function and the illocutive function are referred to here as "linking". Linking and hence the semantic concept of finiteness expresses that a particular state of affairs is claimed to be true for the spatio-temporal anchorpoint talked about. Claiming a state of affairs to be true is the expression of the illocutive function of assertion. Particular linking devices are used to express other illocutive functions such as that of a wish to come true, a permission, or a claim that the achievement of some state of affairs is someone's ability or someone's obligation.

Both the anchoring and illocutive function contribute to the semantic concept of finiteness, each in their own way. Even if no explicit temporal anchoring is provided as in example (1), the form of the finite verb tells us that we talk about a time in the past. But we have to infer from the context, which time it exactly is. Thus, morphological finiteness *reflects* features of the anchoring-information. On the other hand, as shown in (2) finiteness may provide the illocutive function of a claim all by itself.

Both functions are fused in one finite verb form in Dutch, but as we will show in section 2, child and adult language learners clearly separate the two. Initially, they express the illocutive function only, that is, the elements used for that purpose (e.g. modal phrases, adverbials and particles) do not adapt

their form to the anchoring information. As we will see, inflected verb forms also occur. Their morphology however, is not used productively. Thus, anchoring as it is established through the morpho-syntactic expression of tense/aspect or agreement does not occur.

Example (3) from a second language learner of Dutch illustrates the point in case. Here *wel* 'indeed' is used to indicate that a particular claim is made, while anchoring with respect to topic time or some external argument obtains no explicit formal expression.

(3) ik wél hard rijen (Ma 2.9)
I indeed fast drive

Only later in the acquisition process, i.e. after the acquisition of auxiliary verbs, this is done with the adequate target language means of morpho-syntactic finiteness which allow the learner to express the spatio-temporal anchoring information, too.

As shown in (3), the first linking expressions do not stem from the category of verbs. Lexical verbs are of course used in the part of the utterance that refers to the situation expressed (*hard rijen* in (3)), but they are not made finite in order to make them also suitable as linking-expressions. At that stage of acquisition, purely functional carriers of finiteness (e.g. auxiliaries) are absent, too. Learners avoid the complex verbal morphology (that not only involves markings relevant to the semantic concept of finiteness, but also formal adaptations like person and number agreement) and prefer lexical solutions when linking needs to be marked explicitly. It is only at later stages of acquisition that learners develop linking elements that can carry the specific combination of information relevant to the semantic concept of finiteness in one form, namely finite verbs. At the same time, learners must work on the devices for context embedding as well, since the form of the finite verb depends on the kind of (temporal) anchoring chosen.

In what follows, we will take the *illocutive function of finiteness* as a starting point and investigate which kinds of linking devices are used by different types of learners (namely children acquiring Dutch as their L1 and adults acquiring Dutch as an L2) in order to express the semantic concept of finiteness.

The paper is organized as follows. We will first give a short overview on the data sources from first and second language learners (section 1). In section 2 we present a stage model for the development of illocutionary force. Assuming that the structure of early learner languages is based on principles of information structuring an analysis of constituents in terms of

their topic, predicate and linking function seems appropriate. Within this model it is the linking device which expresses the semantic function of finiteness. We will therefore show developmental evidence of the acquisition of finiteness as a grammatical linking device. Section 3 contains a summary of the main points and some conclusions.

2. DATA SOURCES

The data originate from studies on the acquisition of Dutch by children learning this language as their native language and adults learning this language in an untutored second language learning environment. The L1 and L2 data were collected longitudinally. The L1 data are taken from an extensive diary study and cover the period indicated in (4). The L2 data stem from the European project "Second Language Acquisition by Adult Immigrants", funded by the European Science Foundation. In that project, data collection was organized in three cycles of 10 months each, which means that similar types of elicitation techniques (e.g. film retellings and picture descriptions) were repeated three times during the process of data collection (see Perdue 1993). In the present study, we will investigate data from four learners, two with L1 Moroccan (MA) and two with L1 Turkish (T). In the examples below a reference such as, for example, 'Jasmijn 1;9' means that the utterance occurred when Jasmijn was 1 year and 9 months of age, 'Mahmut 1.2' means that the utterance by Mahmut occurred at month 2 of cycle 1.

(4) Data sources. MA = L1 Moroccan Arabic, T = L1 Turkish

L1 Dutch	diary data	J = Jasmijn (1;7-2;2) A = Andrea (1;8-2;4)
L2 Dutch	transcripts of ESF audio recordings	Mo = Mohamed (MA) F = Fatima (MA) Ma = Mahmut (T) A = Abdullah (T)

3. THREE STAGES IN THE DEVELOPMENT OF ILLOCUTIONARY FORCE

As outlined in the previous section we will choose the perspective of information distribution for our analysis of utterance structure. This analysis provides insight into language learning as a developmentally staged process.

Although stages of acquisition are part of a developmental process in which one stage is eventually given up in favour of another, in the actual process of language acquisition stages may overlap, i.e. properties of different stages may co-occur. For some time learner languages may even show evidence of backsliding when learners regress to a stage that they had seemed to have given up. Stages in language development are defined with respect to coherent systems of linguistic structure used by learners of a particular target -language for at least a certain period of time. Our analysis of developmental progress provides evidence that it seems possible in both the L1 and the L2 learning contexts to identify learner varieties at three levels of acquisition.

In the following we will give a description of three types of learner languages used at consecutive stages of linguistic development. These developmental stages are termed: the '*Holistic Stage*', the '*Conceptual Ordering Stage*' and the '*Finite Linking Stage*'. A common property at each of these stages of acquisition is the expression of a topic (explicitly or implicitly) and a state of affairs, such that the state of affairs is claimed to hold for the topic. This relation is established through what we call a validation or linking device. It is this relation of linking which is realised by different linguistic expressions at consecutive stages of language acquisition. At the Holistic Stage validation is achieved by pragmatic means, at the Conceptual Ordering Stage by lexical means and at the Finite Linking Stage by morpho-syntactic means. Interestingly, however, these validation devices are remarkably similar across the L1 and L2 language contexts we investigated.

3.1. The Holistic Stage: Pragmatic Validation

At the Holistic Stage topic and predicate (TOPIC, PRED) are used in a juxtaposition relation. Most frequently no illocutionary element is expressed at all. In these cases, the intended relation between predicate and topic has to be inferred from the discourse context in which the utterance occurs. See, for example, (5a/b) below. A small class of illocutionary phrases (ILP) can occur in either initial or final position. These illocutionary phrases function holistically, i.e. as clausal operators. See (5c-f) and (6a-g) below.

(5) The Holistic Stage in L1 Dutch. Utterances by Jasmijn (1;7-1;9) and Andrea (1;9-1;11)

	ILP	TOPIC TOPIC	PRED PRED	ILP
(a)		poessie *kitty*	jassie bijte *coat bite* (J 1;8)	
(b)		hier *here*	poessie krabd *kitty scratched* (J 1;9)	
(c)	nee *no*	Cynthia *C*	afpakke *away-snatch* (J 1;9)	
(d)		pop *doll*	pot *kaput*	nee *no* (A 1;9)
(e)	unne *want*	Mijnie *M*	dit sijfe *this write* (J 1;7)	
(f)		Cynthia *C*	teenie *toe*	magnie *may-not* (J 1;8)

(6) The Holistic Stage in L2 Dutch. Utterances by Fatima (F), Mahmut (Ma), Abdullah (A), and Mohammed (Mo)

	ILP	TOPIC TOPIC	PRED PRED	ILP
(a)	niet *no*	mijn familie *my family*	buitenland *foreign country* (Ma 2.5)	
(b)	ja *yes*	die meisje *the girl*	brood weggooien *bread away-throw* (Ma 2.9)	
(c)	nee *no*	hier *here*	huis *house* (Ma 1.2)	
(d)			veel eten *much eat*	nee *no* (F 1.5)
(e)	ookwil *too want*	altijd *always*	naar kerk *to church* (F 1.8)	
(f)	misschien *maybe*	ik *I*	beter *better* (F 2.4)	
(g)		ik daar *I there*	vakantie *vacation*	hoefniet *has-to-not* (Ma 2.5)

Examples show that utterances produced by children learning Dutch as their mother-tongue and utterances produced by adult learners of Dutch are

strikingly similar. Topic and predicate are used in a juxtaposition relation, while a small class of holistic illocutionary phrases functions as clausal operators either in initial or final position. Predicates are commonly realised by a nominal element, an adverb, an adjective or a verb particle rather than by a verb (see Herr-Israel & McCune, this volume). If modal operators are used as in (5c-f) and (6a-g) they are precursors in the acquisition of the functional properties of finiteness.

In early child Dutch *nee* 'no' most prominently functions as a clausal operator with a negative modal meaning (5c/d). The same use of *nee* in Dutch can be observed in adult second-language learners (6c/d). With respect to its form as well as its distribution holistic *nee* in early learner varieties of Dutch is modelled after anaphoric *nee*, which is used as an answer to a yes/no-question. It explains why learners may initially use *nee* with scope over the entire clause structure. The clause in the scope of *nee* serves in fact as the explicit expression of a presumed yes/no-question. A similar finding has been reported in Bloom's (1970) study on the use of 'no' in child English. Here, too, 'no' is interpreted as a discourse-semantic means to express non-existence, rejection or denial.

Since the modal operator *nee* has scope over the clause structure as a whole, it is referred to as 'holistic *nee*'. Note, however, that *nee* is used mostly in utterances without a topic. Thus, unambiguous evidence of the holistic use of *nee* is not that frequent. Nevertheless, the holistic use of the negator is productive and is therefore considered a systematic pattern at the initial stage of acquisition.

Simultaneously with holistic *nee* in her early child Dutch, Jasmijn (1;7-1;9) also uses the expressions *unne, minne, ninne, hunne, ulle* 'want' etc. as a positive counterpart (5e). These holistic *unne*-forms seem to be modelled after the adult verb *willen* 'want'. They are productively used in precisely the same position in which *nee* occurs. They function as the positive alternative to *nee* both with respect to their distribution as to their meaning. Given that the meaning of *nee* and *unne* etc. can be described as '[I] do not want' and '[I] want', both devices serve to express 'volition'. Similar types of utterances can be found in L2 Dutch. As the positive alternative to *nee* (6c/d) Fatima uses *ook wil* 'also want' (6e). It shows that she uses both *nee* and *ook wil* to express 'volition'.

Andrea (2;0-2;1) uses *nee* systematically with the same distribution as *mag-ikke* 'may-I'. At the relevant stage, *mag-ikke* is used as a fixed phrase incorporating the topic element *ikke*. Examples of the distributional opposition between *nee* and *mag-ikke* are given in (7).

(7) Holistic use of '*nee'* (no) and '*mag-ikke'* (may-I) in Andrea (2;0-2;1)

(a)	nee	kijke
	no	look (A 2;0)
(b)	nee	bad zitte
	no	bath sit (A 2;1)
(c)	mag-ikke	ook gijbaan?
	may-I	also slide? (A 2;0)
(d)	mag-ikke	ijssie hebbe
	may-I	ice-cream have? (A 2;1)

From this opposition it seems that Andrea uses both *mag-ikke* and *nee* to express 'volition'.

In both child and adult learners elements such as *nee* 'no', *unne* 'want', *ookwil* 'too-want', *mag-ikke* 'may-I' are used to express 'volition'. These illocutionary phrases are functioning with scope over the clause structure as a whole. In the same position learners may also use elements with other kinds of modal meanings. Thus, *misschien* 'possibly' and *magnie* 'may-not' are used to express presence or absence of 'permission', while *hoefnie* 'has-to-not' indicatesabsence of 'obligation'. In the examples above, modal elements are used as means to express illocutionary force. Absence of modality seems to be the default case. Illocutionary phrases used to express absence of modality are the holistic use of *ja* 'yes' and *niet* 'not'. Their illocutive meaning is 'assertion' or 'denial'.

Table 1 summarises the list of illocutive phrases and pragmatic functions at the Holistic Stage that both child and adult learners of Dutch are able to express. The particular ILPs they use, typically occur holistically, i.e. with scope over the predicate as a whole.

Table 1. ILPs at the Holistic Stage

illocutive function	*illocutionary phrase in initial/final position*
assertion	*ja; niet*
volition	*ulle, mag-ikke, ookwil; nee*
ability	*kanwel; kanniet*
permission	*misschien; magnie*
obligation	*hoefnie*

3.2. The Conceptual Ordering Stage: lexical validation

The term 'conceptual ordering' refers to the fact that both the selection and the sequential ordering of constituents in learner grammar is determined by principles of information structuring. At the relevant stage of acquisition learner utterances consist of three structural positions each for constituents with a particular informational function. The 'topic' occurs in initial position. It functions as an anchoring element, i.e. it establishes external reference to the outside world or to the previous utterance. The predicate occurs in final position. It refers to a particular state of affairs which holds for the topic element. The relation between the predicate and the topic element is established by a linking element (ILP). These linking elements occur between the topic and the predicate. Syntactically, they function as adjuncts, i.e. there is no relation of categorical selection between the linking element and the predicate, nor does the linking element entertain a formal relation with other constituents. Given the absence of functional categories, the ordering of the structural positions of topic, linking element and predicate is merely sequential.

Examples of L1 and L2 Dutch utterances at the Conceptual Ordering Stage are given in (8) and (9).

(8) L1 Dutch utterances at the COS (Jasmijn 1;10-1;11; Andrea 2;0-2;1)

	TOPIC	ILP	PRED
(a)	Mijnie *M*	kan *can*	losmake *loose-make* (J 1;10)
(b)	da *there*	kanniet *cannot*	zitte *sit* (A 2;1)
(c)	Jaja *J*	mag *may*	dop opdoen *lid on-do* (A 2;0)
(d)		mánniet *may-not*	doen *do* (A 2;0)
(e)	Peter *P*	moet *has-to*	zitte *sit* (J 1;11)
(f)	disse *this-one*	hoeniet *has-to-not*	meeneme *with-take* (A 2;1)

(g)		doetie *does-he*	alles opete *all up-eat* (J 1;11)
(h)	dit *this*	nee *no*	afdoen *off-do* (J 1;10)
(i)	poppie *doll*	niet *not*	Jaja help *J help* (A 2;0)
(j)		gaag *w'd-like*	boekje leze *book read* (A 2;0)
(k)	papa *daddy*	eve *just*	make *fix* (A 2;1)
(l)	Ruti *R*	hel *indeed*	bad zitte *bath sit* (A 2;1)
(m)	ikke *I*	ook *too*	boot hees *boat been* (A 2;0)
(n)	poesje *kitty*	0	vinger bijte *finger bite* (J 1;10)

(9) L2 Dutch utterances at the COS (Fatima, Mahmut, Abdullah)

	TOPIC	ILP	PRED
(a)	ik altijd *I always*	wil *want*	zit met nabil *sit with N* (F 3.2)
(b)	eenduizend *1000*	kanwel *can-indeed*	sparen per jaar *save per year* (Ma 2.2)
(c)	ik *I*	kanniet *cannot*	praten nederlands *speak Dutch* (F 1.8)
(d)	dan *then*	magniet *may-not*	rijen *drive* (A 2.7)
(e)	vrouw *woman*	moet *must*	keuken *kitchen* (Ma 1.7)
(f)	nog drie maand ik *yet three month I*	moet *have-to*	trouwen *marry* (A 1.6)
(g)	ik *I*	hoefniet *have-to-not*	ziektewet *health-insurance* (Ma 2.7)

(h)	ik *I*	ben *am*	moet werken *have-to work* (A 1.3)
(i)	hij *he*	is *is*	liegt *lies* (A 1.6)
(j)	turkse soep *Turkish soup*	ja *yes*	goed *good* (Ma 1.7)
(k)	dames *ladies*	niet *not*	goeje chauffeur *good driver* (Ma 2.9)
(l)	die jongen *that boy*	misschien *maybe*	weten niet *know not* (A 1.7)
(m)	dan auto *then car*	bijna *nearly*	vallen *fall* (Ma 2.9)
(n)	hij *he*	even *just*	zoeken *look for* (A 2.9)
(o)	ikke *I*	wel *indeed*	handwerk *handwork* (Ma 2.4)
(p)	vrouw *woman*	ook *also*	veel kijk van man *much look to man* (F 1.6)
(q)	viertwintig juni mij 0 man *24 June my* *husband*		thuis *at home* (F 1.8)

As can be seen in (8) and (9) spatio-temporal and deictical elements such as adverbs, proper names and pronouns occur in initial topic position. *The topic* is the constituent that the utterance is about, and it places the utterance in a particular context. From an information-structural perspective therefore, the topic element has *anchoring function*. This explains why at the relevant stage children have no *wh*-question words in topic position. The topic position may be left empty if reference can be inferred from context.

Constituents in final position contain *non-finite, VP-like elements*. They refer to a particular state of affairs which holds for the element in topic position. From an information-structural perspective therefore, these VP-like elements have *predicate function*.

As pointed out, the relation between the predicate and the topic element is established by a *linking element*. This linking element occurs between the topic and the predicate. Linking devices at the Conceptual Ordering Stage are morphologically fixed expressions that are used to validate relations between

the predicate and the topic. It is their linguistic function to express properties of *illocutionary force*. They are referred to as illocutionary phrases (ILP).

ILPs are *proto-functional elements* with *pragmatic properties* of finiteness. They constitute a closed class of lexical phrases containing:

1. modal verb-like elements: *unne* 'want', *wil* 'want', *kanniet* 'cannot', *magwel* 'may-indeed', *moettie* 'has-to-it', *hoenie* 'has-to-not'
2. light verbs: *doetie* 'does-it', *gaatie* 'goes-it', *ben* 'am', *is* 'is'
3. adverb-like elements: *nee* 'no', *ja* 'yes', *niet* 'not', *handigniet* 'handy-not'
4. modal particles: *g(r)aag* 'want-please', *eve* 'want-just'
5. scope particles: *wel* 'indeed', *ook* 'too', *zelf* 'myself'
6. empty positions

Despite their difference in origin, these modal operators are all used with the same distributional properties. That is elements such as *wel* 'indeed', *ook* 'too' and *nee* 'no' are used the same way as *magwel* 'may-indeed' or *kanniet* 'cannot'. From this it seems fair to conclude that these linking elements are lexical phrases not (yet) to be categorised as expressions of a verbal category. This explains why at the relevant stage target-like finite verb forms among them auxiliary verbs are systematically absent.

The linking elements constitute a syntactic category which is especially relevant for early child and adult second language learner varieties. They can be analysed as operators used to express 'volition', 'ability', 'permission', 'obligation' and 'assertion'. While modality expresses different kinds of illocutionary force, the default value of assertion is established by absence of modality. Positive assertion can therefore be expressed by leaving the position of the linking device empty.

Differences in contrast are signalled via intonation. Dimroth (2002) argues that a stressed particle with scope to the left, such as stressed *auch* 'too' in German, is used to indicate contrastive topic. Similarly in child L1 and adult L2 Dutch, *ook* 'too' is used to indicate that a particular state of affairs is not only relevant for the topic-referent, but *for another or other referents,* too. Use of *zelf* 'self' indicates that a particular state of affairs holds for the topic-referent and *not for any other referent*. Finally, *wel* 'indeed' is used if the state of affairs holds for the topic-referent but *not for a particular other referent*. Table 2 summarises the illocutive phrases and pragmatic functions at the Conceptual Ordering Stage that both child and adult learners of Dutch are able to express.

Table 2. ILPs at the Conceptual Ordering Stage

illocutive function	*illocutive phrase*	*scope particle (+ contrast)*
assertion	*0, doetie*, *gaatie; niet*	*0*, *ook, wél*
volition	*ulle, ja*, *mag-ikke; nee*; *graag*, *eve*	*óók, zélf, wel*
ability	*kanwel; kanniet, handigniet*	*ook*, *zelf, wel*
permission	*magwel; magniet*	*ook, wel*
obligation	*moettie; moetniet, hoefniet, niet*	

As its name indicates, the topological structure of the Conceptual Ordering Stage mirrors the basic information-structure of the utterance. Although the Holistic Stage showed a preference towards a topic-predicate order, too, it nevertheless did not allow for a systematic expression of the relation occurring between the element(s) talked about, the topic, and what is said about them, the predicate. At the Conceptual Ordering Stage it is now possible for the learner to differentiate between various forms of the relation between topic and predicate and, moreover, to express conditions on it by use of scope particles as linking elements. Thus, the most important step forward consists here of a lexicalised link between the topical element(s) and the predicate. By way of illustration the following examples are taken from (5), (6) and (8), (9).

(5a) poessie jassie bijte
kitty *coat bite* (J 1;8)

(6a) niet mijn familie buitenland
not my family *foreign country* (Ma 2.5)

(8a) Mijnie kan losmake
M *can* *loose-make* (J 1;10)

(9a) ik altijd wil zit met Nabil
I always *want* *sit with N* (F 3.2)

Utterances as in (5a) and (6a) are typical of the Holistic Stage. At the relevant stage there is no element to express the relation between the topic and the predicate. Utterances as in (8a) and (9a) are representative of the Conceptual Ordering Stage. Here, a particular element systematically occurs

in linking position. Given its illocutive function it validates the relation between the topic and the predicate.

3.3. Finite linking stage: grammatical validation

The transition from the Conceptual Ordering Stage to the Finite Linking Stage is characterised as a transition from lexical to grammatical validation of the utterance. At the Finite Linking Stage, finite verbal morphology is the typical reflex of the validation operation. Validation is obligatorily expressed through finiteness marking. It occurs via morphological adaptation of the linking element to features of the topical elements (tense, person/number) and placement in a (syntactic) operator position with respect to the predicate. In addition to that, the validation device sometimes gets fused with parts of the predicate, namely the lexical verb. In spontaneous production this fusion seems to correlate with the Aktionsart properties of the verb.

For child and adult learners of Dutch, this transition poses three distinct acquisitional problems. Firstly, finite utterances in the target languages are marked by verb morphology, thus the 'link' in non-elliptic utterances is always explicit. The learner must come to realise that this obligatory verb morphology carries both temporal/aspectual and person/number agreement information. Secondly, the finite marking falls on the leftmost verbal item, a fact which may lead to a merge between the link and the verbal part of the predicate. Thirdly, learners have to acquire Verb-second which means that in declarative sentences there is only one constituent position before the finite verb.

In both children and adults we see the acquisitional task accomplished, firstly, by the acquisition of the auxiliary system. While the 'conceptual' ordering is maintained, the learners work on the 'link' as marked by modal and auxiliary verbs. As will be shown in 3.3.1 the acquisition of the auxiliary system entails a syntactic reorganisation in terms of a functional category system. It establishes both a head-complement relation between the auxiliary and the lexical verb and a relation of morphological agreement between the auxiliary verb and the external argument. Finite lexical verbs become productive only later. Finally, the acquisition of Verb-second involves a choice to be made between items competing for a one-constituent topic position. As will be shown in 3.3.2 this phenomenon takes time to master.

3.3.1. The acquisition of the auxiliary system

In early child Dutch there is a drastical increase in the use of modal and auxiliary forms at the Finite Linking Stage. However, the development of auxiliaries with past participles lags behind that of modals with infinitives: whereas the latter are productive at the Conceptual Ordering Stage, the former only become productive at the Finite Linking Stage. The relevant figures for Jasmijn and Andrea (Jordens 2002:745) are given in Table 3.

Table 3. The acquisition of modal and auxiliary verbs in L1 Dutch

Child	*age*	*inf.*	*'doet'/'gaat'/ MOD + inf.*	*%*	*pp.*	*'heeft'/'is' + pp.*	*%*
Jasmijn	1;10-1;11	243	107	30%	55	2	4%
	2;0-2;2	85	134	61%	30	18	38%
Andrea	2;0-2;1	295	61	17%	29	3	9%
	2;2-2;4	158	161	50%	32	53	62%

In Jordens (2002) it is argued that the development of the category of auxiliaries (*hebben/zijn*) + a past participle (pp) can be seen as a bootstrap into the grammatical aspect system of the target language system. Concomitantly, *gaan* 'go' and *doen* 'do' + infinitive are developed to express aspectual values other than the perfect. These four auxiliary verbs are the first to show person/number agreement. The development of a full pronoun syst-em accompanies the development of inflectional agreement with modal verbs too. Examples of this can be seen in (10).

(10) The acquisition of auxiliary and modal verbs to express aspect in L1 Dutch

	TOPIC	AUX	VP
(a)		heef	Cynthia maakt
		has	*C made* (Jasmijn 2;0)
(b)	ik	heef	óók appel gete
	I	*have*	*too apple eaten* (Andrea 2;4)
(c)	da	ben	ikke ook wees
	there	*am*	*I also been* (Andrea 2;3)

(d)	die	is	altijd opde televisie geweest
	that-one	*is*	*always on TV been* (Jasmijn 2;2)
(e)	ik	doet	neusje snuite
	I	*do*	*nose blow* (Jasmijn 2;0)
(f)		doen	ze *same* hope
		do	*they together walk* (Andrea 2;2)
(g)	ik	ga	eve die glije
	I	*go*	*just that-one slide* (Jasmijn 2;0)
(h)		gaan	ze almaal ete, zie?
		go	*they all eat, see?* (Andrea 2;3)
(i)	Jaja	wilt	óók dakik ijsje opete
	J	*wants*	*too soon ice-cream up-eat* (Andrea 2;3)
(j)	daa	kan	ik niet meer lope
	there	*can*	*I not anymore walk* (Jas. 2;2)
(k)	die	mag	je nog hebbe
	that-one	*may*	*you also have* (Jas. 2;2)
(l)	dan	moet	Cynthia weer make
	then	*has-to*	*C again make* (Jasmijn 2;2)

Auxiliary verbs serve as a bootstrap in the transition to morpho-syntactic properties of finiteness. This becomes evident from a comparison of the examples in (8) at the Conceptual Ordering Stage with the examples in (10) at the Finite Linking Stage. With the acquisition of auxiliary verbs the class of elements in linking position as well as their morpho-syntactic properties haschanged significantly:

1. The unanalysed phrasal expressions with an illocutive meaning have disappeared. There are no more phrases such as *unne* 'want', *kanniet* 'cannot' vs. *kanwel* 'can-indeed', *magniet* 'may-not' vs. *magwel* 'may-indeed', *mag-ikke* 'may-I', *doetie* 'does-he/it', *gaatie* 'goes-he/it';
2. Adverb-like elements 'modal particles' such as *nee* 'no', *handigniet* 'handy-not', *niet* 'not', *g(r)aag* 'please', *eve* 'just' do not occur in linking position;
3. Scope particles such as *wel* 'indeed', *ook* 'too', *zelf* 'self' do not occur in linking position;
4. The linking position cannot be empty;

5. In linking position there is a closed class of auxiliary and modal verbs with target-like morphology: *heb/heeft* 'have/has', *ben/is* 'am/is', *mag* 'may', *kan* 'can' etc.

Developmental processes as in 1 to 5 are evidence that finiteness as a lexical-semantic category has been given up in favor of finiteness as a functional verb-category. Furthermore, within the category AUX, there is a distributional opposition between *heeft,is /doet /gaat*. Examples are given in (11) and (12). Particularly, Jasmijn's (2;0-2;2) examples with the lexical verb *maken* 'make' and Andrea's (2;2-2;4) examples with the lexical verb *eten* 'eat' show that these auxiliaries are used in complementary distribution. They function as periphrastic means to express aspectual distinctions.

(11) The use of Auxiliary verbs in Jasmijn (2;0-2;2) to express grammatical aspect

perfect: *heb* 'have-SG', *heef(t*) 'has'; *ben* 'am', *is* 'is', *was* 'was'
heef Cynthia maakt (2;0) / (has C made)
ik *was* poepie doen (2;0) / (I was poup do)
ik *heef* afspoeld (2;2) / (I have washed)
die *is* altijd opde televisie geweest (2;2) / (that-one is always on TV been)
waar *ben* je nou geweest? (2;2) / (where are you now been?)
die *heb* ik wel geplakt (2;2) / (that have I glued)

imperfective: *doe* 'do-SG', *doet* 'does'
doe je Pino make? (2;0) / (do you P make?)
ik *doet* neusje snuite (2;0) / (I do nose blow)
poesje *doet* likke (2;1) / (cat does lick)
doet mama mij ... (2;2) / (does mommy me ...)

perfective: *ga* 'go-SG', *gaat* 'goes'
ik *gaat* Pino make (2;0) / (I go P make)
ik *ga* eve die glije (2;0) / (I go just that slide)
ga je ook school toegaan? (2;1) / (go you too school to-go)
gaat Cynthia slape? (2;1) / (goes C sleep?)

(12) The use of Auxiliary verbs in Andrea (2;2-2;4) to express grammatical aspect

perfect: *hemme* 'have-SG', *heef* 'have-SG'; *ben* 'am', *is* 'is'
ikke *hemme* deze tekend (2;3) / (I have this drawn)
da *ben* ikke ook wees (2;3) / (there am I also been)
isse barnies affehope mam? (2;4) / (is 'barnies' finished mommy?)
ik *heef* óók appel gete (2;4) / (I have too apple eaten)

imperfective: *doe* 'do-SG', *doet* 'does', *doen* 'do-PL'
doen ze same hope (hope = lopen) (2;2) / (do they together walk)
Jaja *doet* kitkat opete (2;3) / (J does kitkat up-eat)
ik *doe* mij fesje aan mij jamaboek doen (2;4) / (I do my vest on my pyjamas do)

perfective: *gaat* 'goes', *gaan* 'go-PL'
Jaja *gaat* daar zitte en papa *gaat* daar zitte (2;2) / (J goes there sit and daddy goes there sit)
gaan ze almaal ete, zie? (2;3) / (go they all eat, see?)
gaat ikke ook mee naa paardrije? (2;4) / (go I also with to horse-ride)

As soon as the auxiliaries *heb/heeft*, *gaa(t)* and *doe(t)* have come to be used to express aspectual distinctions, the relation between the elements in second position and the predicate has been reinterpreted, too. When children have established a grammatical relation between auxiliary verbs, on the one hand, and the predicate on the other, they have, in fact, discovered the relation between the structural position of a head and its complement. Whereas the illocutionary elements of the Conceptual Ordering Stage are adjuncts, auxiliary verbs are part of a functional category system. As is the case in the target language, they function as the head of a head-complement structure. The fact that auxiliaries are part of the functional category system of the target language, i.e. the fact they have no lexical meaning, explains why they do not occur at the Conceptual Ordering Stage. The relevant stage is referred to as the Finite Linking Stage, because it is characterised by the fact that morpho-syntactic features are used to establish structural relations between the Topic and the Predicate.

At the Finite Linking Stage modal verbs (elements of the category MOD) are to be seen as particular instantiations of AUX. They carry the grammatical properties of prospective aspect, while their lexical meaning is used to express illocutionary force, i.e. 'volition', 'ability', 'permission', and

'obligation'. As in the case of the Conceptual Ordering Stage, absence of modality implies the expression of assertion. At the Finite Linking Stage, it is the auxiliaries that are used to carry the illocutionary function of 'assertion'. The functional system of the linking elements at the Finite Linking Stage is given in Table 4.

Table 4. The auxiliary system to express aspect

ASP	AUX		ILF
perfect:	*heb*, *heeft*; *ben*, *is*		assertion
imperfective:	*doe(t),*		assertion
perfective*:*	*ga(at)*		assertion
prospective:	MOD:	*wil, wilt*	volition
		kan	ability
		mag	possibility
		moet	obligation

At the Conceptual Ordering Stage modal and scope particles occurred in complementary distribution with modal phrases. At the Finite Linking Stage, they have not disappeared nor have their scope properties. However, while they were used in complementary distribution with modal phrases at the Conceptual Ordering Stage, they can no longer be used without auxiliaries or modal verbs at the Finite Linking Stage.

Summarising, the acquisition of auxiliaries leads to a reanalysis of the illocutionary elements from the Conceptual Ordering Stage as grammatical elements with an aspectual function at the Finite Linking Stage. Those lexical-semantic and grammatical-syntactic properties which are relevant to this process of restructuring are given in Table 5.

Table 5. From 'Illocutionary force' to the function of a 'Head'

	Conceptual Ordering Stage: proto-MOD	Finite Linking Stage: AUX
grammatical status:	lexical category	functional category
meaning:	illocutionary force	grammatical aspect
syntactic function:	adjunct	head

At the same time, the acquisition of morphological person/number agreement with an external argument is evidence of the acquisition of a specifier-head relation between the NP and the auxiliary. Having acquired

both the head-complement and the specifier-head relation with auxiliary verbs the children have learned the grammatical properties of the functional category system which determines the projection of the inflectional phrase in Dutch. Thus, the acquisition of both the head-complement and specifier-head relationship can be traced to the fact that the auxiliary verbs *hebben* and *zijn* serve as a bootstrap into the functional category system of the target language.

With the adult learners too, it is only from the grammaticalized finiteness phase on that formal agreement with contextual information appears in the linking slot, in the form of morphological marking, first on the auxiliary, and ultimately on lexical verbs. Examples are given in (13).

(13) Finiteness marking on non-thematic verbs in L2 Dutch: Fatima (F), Mahmut (Ma), Abdullah (A), and Mohamed (Mo)

	TOPIC	AUX	VP
(a)	ik	heb	alleen spaar die geld van die winkel
	I	*have*	*only save(d) the money from the shop* (F 3.8)
(b)	ik	heb	niet zeggen
	I	*have*	*not say* (A 2.7)
(c)	ik	ben	net gezegd
	I	*have*	*just said* (A 1.8)
(d)	jij	is	voor mij betaald
	you	*is*	*for me paid* (F 2.6)
(e)	hij	is	slaap
	he	*is*	*sleep* (F 3.6)
(f)	nu auto	is	afgelopen
	now car	*is*	*run-out* (F 3.9)
(g)	die	ben	getrouwd
	there	*am*	*married* (Ma 1.5)
(h)	ik	doe	niet open
	I	*do*	*not open* (A 3.3)
(i)		doet	kijk
		does	*look* (F 2.7)
(j)	ik	ga	werk zoeken
	I	*go*	*work look-for* (A 2.1)

(k)	un klein flat,		
	a small flat,		
	die	wil	ik niet
	that	*want*	*I not* (Mo 3.6)
(l)	hier	kan	ik niet lezen, hè
	here	*can*	*I not read, ok* (Mo 2.8)
(m)	ik	moet	nog opereerd
	I	*must*	*still be operated* (Mo 3.4)

As shown in (13b,h,k,l) the negator is placed to the right of the auxiliary verbs and to the left of lexical verbs (see also Parodi 2000, and Becker to appear, for a discussion of similar phenomena in L2 German). Particles and adverbs, such as *ook* 'also' and *nog* 'still' (13m), are progressively integrated into the utterance structure, starting to appear behind auxiliary verbs. Their domain of application is no longer necessarily adjacent. In *dan die meisje ook slapen* 'then the girl also sleep' *ook* is right-adjacent to its domain of application *die meisje* 'the girl', whereas in *dan die meisje is ook slapen* 'then the girl is also sleep' the same domain of application would be non-adjacent.

Normally developing children are able to achieve this stage of language development more or less at the same period of time. However, not all adults are able to accomplish the same. Thus, while the examples of Fatima and Mahmut in (13) are all the utterances of this kind occurring in their data, the examples of Abdullah and Mohammed in (13) are only a few among many others. As with children in L1 Dutch, Abdullah and Mohammed use auxiliary verbs in distributional opposition. Examples are given in (14) and (15). Illustrative are Abdullahs examples with the lexical verb *werken* 'work' and Mohammeds examples with the lexical verbs *slapen* 'sleep', *lopen* 'walk', *vallen* 'fall', *staan* 'stand'. They show that the auxiliaries are used in complementary distribution. They function as periphrastic means to express aspectual distinctions.

(14) The use of auxiliary verbs in Abdullah to express grammatical aspect

perfect: *heb* 'have-SG', *heeft* 'has'; *ben* 'am'
ik *heb* werk gevonden (A 3.5) / (I have work found)
hij *heeft* niet gezien (A 1.9) / (he has not seen)
die jongen *heeft* gewonnen (A 3.3) / (that boy has won)

ik *ben* alleen werkt elfhonderd gulden (A 3.3) / (I am alone worked elevenhundred guilders)

imperfective: *doen* 'do-SG/PL'
wij *doen* die nylon draaien (A 3.5) / (we do that nylon turn)
soms *doen* ik overwerken (A 3.6) / (sometimes do I overwork)

perfective: *ga* 'go-SG', *gaat* 'goes'
ik *ga* werk zoeken (A 2.1) / (I go work look-for)
dan hij *gaat* hier zitten (A 2.9) / (then he goes here sit)
ik *gaat* uitzendbureau werken (A3.7) / (I goes secretarial bureau work)

(15) The use of auxiliary verbs in Mohamed to express grammatical aspect

perfect: *heb* 'have', *heeft* 'has'; *ben* 'am', *is* 'is', *was* 'was'
hij *heeft* hem niet gezien (Mo 2.6) / (he has him not seen)
ik *heb* niet geslapen (Mo 3.1) / (I have not slept)
ik *heb* die marokkaan geslaan (Mo 2.8) / (I have that moroccan beaten)
ben daar nooit geweest (Mo 3.7) / (am there never been)
toen hij *is* liept (Mo 3.9) / (then he is walked)
die politie *was* daar vallen (Mo 3.9) / (that police was there fallen)

imperfective: *doe* 'do-SG', *was* 'was', *staat* 'is'
dan ik *doe* boks (Mo 2.8) / (then I do boxing)
hij *was* zo snel lopen, hè (Mo 2.6) / (he was so fast walk)
kinderen *was* staan bij hem (Mo 2.9) / (children was stand with him)
toen ik *was* bij hun slapen (Mo 3.5) / (then I was with them sleep)
toen die politie *staat*ie bellen (Mo 3.9) / (then that police stand-he call)

perfective: *ga* 'go-SG', *gaat* 'goes', *gaan* 'go-PL'
die oma *gaat* bij hem staan (Mo 1.9) / (that grandmother goes with him stand)
dan ik *ga* rustig zitten (Mo 3.9) / (then I go quietly sit)
als iemand doet bij jou zo *gaat* jij vallen, hè? (Mo 2.7) / (if someone does with you so go you fall)
toen die meisje *gaat* snel lopen (Mo 3.9) / (then that girl goes fast walk)
en toen allebei *gaan* slaap (Mo 2.3) / (and then both go sleep)

3.3.2. The acquisition of Verb-second

As the ultimate step to the Finite Linking Stage, children's grammars appear to develop a specifier position within the VP constituent, which was not available at the Conceptual Ordering Stage. As a result of this, the initial position becomes available as a possible landing site for constituents such as the internal argument or an adverbial element, while at the same time the external argument occurs in the specifier position of VP. This syntactic phenomenon of topicalization now accounts for the anchoring function of items in the initial position. However, placement in accordance with the verb-second rule of the target languages takes time to master.

For adult learners, as with the children, mastering finiteness and Verb-second is not a one-step development. The examples in (16) show that two constituents may precede the finite verb:

(16) Finite verbs in V3 position in L2 Dutch

(a) dan politie *wilt* charlie slaan (Mo 2.6)
then police wants Charlie beat

(b) toen die oma *wilt* ook auto rijen (Mo 3.9)
then that grandmother wants also car drive

(c) hier vrouw *is* valt (Mo 3.2)
here woman has fallen

(d) vrijdag ik *heb* niet geslapen tot zes uur (Mo 2.8)
Friday I have not slept till six o'clock

(e) wanneer ik *heb* geslaap? (Mo 3.5)
when I have slept?

(f) wat hij *heeft* gedaan (Mo 3.6)
what he has done

(g) un kruising hij *heeft* daar ongeluk gemaakt (Mo 3.9)
a crossing he has there accident made

To summarise, with the mastery of the Finite Linking Stage for children and for adults, if they get that far, positive assertion is expressed by finiteness. Morphologically, finiteness also marks the link to the topic time, as well as person/number agreement to the grammatical subject.

4. SUMMARY AND CONCLUSION

As a result of an analysis based on an information structural approach, striking similarities were observed in the acquisition of Dutch by children learning their mother tongue and adults learning a second language. A common property observed in both learning situations is the expression of a topic and a state of affairs, such that the state of affairs is claimed to hold for the topic. This relation between the topic and the state of affairs is established through what we have termed a validation or linking device. It is this validation device which is expressed differently at consecutive stages of language acquisition, but where each stage shows cross-acquisitional and cross-subject similarities. These different ways to express validation allowed us to discriminate between a Holistic Stage, a Conceptual Ordering Stage and a Finite Linking Stage. At each of these stages of language acquisition the relation of linking or validation is expressed by different linguistic means. At the Holistic Stage validation is mainly achieved pragmatically, i.e., validation typically depends on variables inherent to the context. At the Conceptual Ordering Stage validation is achieved lexically by the use of elements expressing illocutionary force. At the Finite Linking Stage validation occurs morpho-syntactically by the use of elements with a grammatical function.

The devices of validation turned out to be remarkably similar in children learning their mother tongue and in adults learning a second language. At the Holistic Stage, frequently no validation device is expressed at all. We often find some topic and predicate in a juxtaposition relation and even predicates occurring alone. Characteristic however, is the use of modal operators in clause-initial or clause-final position. These modal operators are used holistically, i.e. they have scope over the topic-predicate structure as a whole. The holistic use of clausal operators occurs most prominently with *nee* 'no' in early child Dutch. It has a negative modal meaning as in 'I don't want'. As positive alternatives Dutch children use elements based on the target verb form *wil(len)* 'want' or *mag-ikke* 'may I'. These modal expressions are used meaning 'I want'. With similar distribution expressions such *ook wil* 'also want' can also be found in adult L2 learners of Dutch.

At the Conceptual Ordering Stage basic utterance structure consists of a sequence of three structural positions each for constituents with a particular informational function. Constituents in these positions are related by adjunction, while their ordering depends on principles of information structuring. Each position can be filled by more than one constituent, which means that target-language syntactic constraints have yet to be acquired. Constituents in initial position function as 'topic' elements. A topic element

can be an NP, often a proper name, or a deictic adverb. Topics establish external reference to the outside world or to the previous utterance. They have 'anchoring function', i.e. they ensure the embedding of the utterance in context. The constituents in end position are termed 'predicates'. Predicates at the Conceptual Ordering Stage are non-finite. They can be a VP or a VP-like expression, referring to a particular state of affairs. The relation between topic and predicate is established by the linking element. It serves as a validation device, i.e. it is used to express that a particular state of affairs expressed by the predicate holds for a particular topic. Validation varies with respect to the kind of relation that is claimed to hold. Modal elements may occur as linking devices in second position. They belong to a closed-class category of modal expressions that are used with illocutionary force. They allow both children and adults to express 'volition', 'ability', 'permission', 'obligation'. In absence of modality lexical linking elements indicate 'assertion'. Since 'assertion' is the default pragmatic function, the position of the element used to express assertion may be left empty. In sum, linking devices constitute (a) elements expressing positive or negative assertion such as *doe(t)* 'does' or *niet* 'not' (b) scope particles such as *wel* 'indeed', *ook* 'also' and *zelf* 'self' and, finally, (c) modal phrases expressing volition, ability, permission, and obligation, such as *mag-ikke* 'may-I', *nee* 'no', *kanwel* 'can-indeed', *kanniet* 'cannot', *magwel* 'may-indeed', *magniet* 'may-not', *moet* 'has-to', *hoefniet* 'has-to-not'. At the Conceptual Ordering Stage, modal phrases are lexical linking devices with no syntactic function.

At the Finite Linking Stage of both children and adult L2 learners validation of the relation between the state of affairs and the topic element is grammaticalized. Acquisition of the auxiliaries *heb*, *heeft*, *ben*, *is* 'have, has, am, is' etc. leads to a restructuring of unanalysed modal expressions. Bootstrapping with members of the target functional category AUX establishes a syntactic category which functions as the head of a head-complement relation. Due to the fact that finite auxiliaries have come to be used as a grammatical linking device functioning as the head of a head-complement structure, particles and zero marking cannot be used any more as a means of validation. The same holds for *nee* 'no' and other target adverbial elements. Furthermore, the morphology of the elements of AUX may also represent the anchoring function of spatio-temporal reference and personal deixis. As such it establishes a specifier-head relation, too. Having acquired both the head-complement and the specifier-head relation with auxiliary verbs children and adults have learned the grammatical properties of the functional category system which determines the projection of the inflectional phrase in Dutch.

Finiteness is a functional property of Germanic languages. It is carried by auxiliaries and lexical verbs occurring in second constituent position (Klein 1998). In order to explain processes of the acquisition of finiteness it seems necessary to distinguish between the semantic concept of finiteness and its formal representation (Klein 1998). The semantic concept of finiteness entails two properties of information structuring: anchoring and illocutionary force. Anchoring is the pragmatic operation which establishes the identification of what is talked about and the embedding of the actual utterance in a discourse world. The illocutionary function of finiteness validates the relation between a particular state of affairs described in the predicate part of an utterance and a topic element. While both functions are fused in one finite verb form in the target languages investigated, they are clearly separated in learner varieties of both children and adult learners.

The standard assumption of the differences in the expression of finiteness between child and adult acquirers is that the former quickly and effortlessly develop the necessary verbal morphology while the latter do not. We have tried to give a more nuanced picture, showing that there are more similarities in the process than is usually assumed.

PETER JORDENS
Vrije Universiteit, Amsterdam

CHRISTINE DIMROTH
MPI for Psycholinguistics, Nijmegen

5. REFERENCES

Becker, Angelika. "The semantic knowledge base for the acquisition of negation and the acquisition of finiteness." In Henriette Hendriks (ed.), *The structure of learner varieties*. Berlin: Mouton de Gruyter, 2004: 349-416.

Bloom, Lois. *Language development: form and function in emerging grammars.* Cambridge, Mass.: The M.I.T. Press, 1970.

Dimroth, Christine. "Topics, assertions and additive words." *Linguistics* 40 (2002): 891-923.

Dimroth, Christine, Petra Gretsch, Peter Jordens, Clive Perdue, and Marianne Starren "Finiteness in Germanic languages: A stage-model for first and second language development." In Christine Dimroth and Marianne Starren (eds.), *Information structure and the dynamics of language acquisition*. Amsterdam: Benjamins, 2003: 65-93.

Herr-Israel, Ellen and Lorraine McCune. "Dynamic event words, motion events and the transition to verb meanings." (this volume).

Jordens, Peter. "Finiteness in early child Dutch." *Linguistics* 40 (2002): 687-765.

Klein, Wolfgang. "Assertion and finiteness." In Norbert Dittmar and Zvi Penner (eds.), *Issues in the Theory of Language Acquisition*. Bern: Lang, 1998: 225-245.

Parodi, Teresa. "Finiteness and verb placement in second language acquisition." *Second Language Research* 16 (2000): 355-381.

Perdue, Clive. *Adult language acquisition: cross-linguistic perspectives*. Cambridge: Cambridge University Press, 1993.

Part 4

Language-specific variation and the role of frequency

DAVID GIL

THE ACQUISITION OF VOICE MORPHOLOGY IN JAKARTA INDONESIAN

Abstract. Whereas in English children typically acquire the passive construction at a relatively advanced age, in some other languages it has been reported that the passive is acquired significantly earlier. This paper examines the acquisition of voice morphology in the Jakarta dialect of Indonesian, drawing on an ongoing longitudinal study of eight children, represented in a corpus of over 500,000 utterances. The paper focuses on the acquisition of two prefixes, typically characterized as markers of passive and active voice, showing that children begin using these prefixes in a productive and appropriate manner at a very early age, sometimes under 2;0. It is argued that the early acquisition of voice morphology in Jakarta Indonesian is due to a combination of two distinct factors: the formal simplicity of the associated morphosyntactic patterns and the high frequency with which the relevant forms occur in the adult input to the process of language acquisition.

1. INTRODUCTION

In English, children typically acquire the passive construction at a relatively advanced age, as shown by Harwood (1959), Brown (1973), Baldie (1976), Horgan (1978), Bridges (1980) and others. Many other languages resemble English in this respect. However, there are also some languages in which the passive voice is acquired at a rather earlier age, for example Zulu (Suzman 1985, 1987), Sesotho (Demuth 1989, 1990), Tagalog (Segalowitz and Galang 1978), Inuktitut (Allen and Crago 1993, 1996) and Quiche (Pye and Quixtan Poz 1988).

Why do languages differ with respect to the acquisition of passive constructions? Since the children learning the different languages are presumably endowed with the same abilities, the variation must lie in the languages themselves: in the inherent properties of the respective constructions, or the ways in which these constructions present themselves to the child.

This paper examines the acquisition of voice morphology in the Jakarta dialect of Indonesian, drawing on an ongoing longitudinal study of eight children, represented in a corpus of over 500,000 utterances. The paper focuses on the acquisition of two prefixes, typically characterized as markers of passive and active voice. And it finds that children begin using these prefixes in a productive and appropriate manner at a very early age, sometimes under 2;0.

N. Gagarina and I. Gülzow (eds.), The Acquisition of Verbs and their Grammar: The Effect of Particular Languages, 201–227.

This paper argues that the early acquisition of voice morphology in Jakarta Indonesian is due to a combination of two distinct factors: the formal simplicity of the associated morphosyntactic patterns and the high frequency with which the relevant forms occur in the adult input to the process of language acquisition.

2. VOICE MORPHOLOGY IN THE ADULT LANGUAGE

Jakarta Indonesian is the general colloquial language used in Jakarta, the capital and largest city of Indonesia, in most everyday contexts, for inter-ethnic and increasingly also intra-ethnic communication; in addition, it is gaining in currency as an informal lingua franca throughout Indonesia. Jakarta Indonesian is acquired naturally, automatically and completely, at a young age, by most or all children growing up in Jakarta. Jakarta Indonesian is quite distinct from Standard Indonesian, used in more formal contexts in Jakarta and throughout Indonesia, and familiar to many general linguists from a large descriptive literature.

This paper is concerned with the two voice prefixes *di-* and *N-*. (The symbol *N-* denotes a process of prenasalization, involving either the prefixation of *ng(e)-* or the replacement of a stem-initial consonant with a homorganic nasal.) A brief sketch of the voice morphology of Jakarta Indonesian is provided in Gil (2002:268-269); the remainder of this subsection fleshes out that sketch in more detail. The main focus of Gil (2002) is on a quite different dialect of Indonesian, that spoken in the province of Riau. However, it turns out that the major syntactic and semantic characteristics of the prefixes *di-* and *N-* are the same in both dialects.

The prefixes *di-* and *N-* are generally considered to be markers of passive and active voice respectively; see for example Mintz (1994), who intersperses data from colloquial Jakarta Indonesian in his description of the standard language. However, as we shall see below, these two prefixes exhibit an array of properties which set them apart from prototypical markers of passive and active voice familiar from other languages. Nevertheless, they still share sufficiently many properties with passive and active markers to warrant their characterization in terms of a broader notion of *generalized voice marker*.

Like the standard language, Jakarta Indonesian has no agreement or case-marking morphology. However, Jakarta Indonesian differs from the standard language in exhibiting a considerably greater degree of flexibility with respect to word order. In a simple transitive clause expressing an activity plus two associated participants, the three elements may occur in any of the 6 logically possible orders, without change in meaning. Thus, Jakarta Indonesian has no grammatical devices, or at least no obligatory ones, for the

differentiation of thematic roles: only real-world knowledge and the specific contexts of individual utterances enable speakers and hearers to assign the appropriate interpretations in each case.[1]

But what of voice morphology, and the prefixes *di-* and *N-*? Several arguments can be adduced suggesting that they should not be analyzed as passive and active markers, at least not in the usual senses of the terms. First, unlike many languages in which passive forms are derived from their active counterparts, in Jakarta Indonesian both *di-* and *N-* forms are derived from a neutral bare stem. Secondly, unlike most languages with passive and active voices, clauses containing *di-* and *N-* forms are of apparently identical syntactic structure. Thirdly, unlike most languages with passive and active voices, the voice marker *di-* is more frequent than its counterpart *N-* (see section 3.2 below).

Whereas the above three arguments carry over also to Standard Indonesian, the following two additional arguments are specific to Jakarta Indonesian. In most languages with passive and active voices, these voices constitute a grammatical paradigm, in the sense that each verbal form must be marked with exactly one voice marker. However, in Jakarta Indonesian, this is not the case. Very often, forms occur bare, without *di-* or *N-*. And occasionally, albeit much less commonly, forms occur with both prefixes, *di-* and *N-*. Thus, for example, for 'attach', with base *tempel*, our corpus contains not only the forms *ditempel* and *nempel* but also *tempel* and *dinempel*: if *ditempel* and *nempel* are labeled as passive and active, then *tempel* would have to be characterized as neither passive nor active, and *dinempel* as both passive and active.[2]

The final argument pertains to the grammatical and semantic function of the prefixes. In most languages with voice markers, one of their core functions is to assign thematic roles to certain participants. For example, in a language such as English, or, for that matter, Standard Indonesian, an NP preceding a passive verb is assigned a thematic role typically associated with objects, such as patient, while an NP occurring before an active verb is assigned a thematic role characteristically associated with subjects, such as actor. In contrast, in Jakarta Indonesian, this function is absent; instead, an analogous pattern manifests itself as a mere statistical tendency. Thus, an NP preceding a word prefixed with *di-* is usually interpreted as the patient, but may occasionally be understood as bearing other thematic roles, including that of actor; similarly, an NP occurring before a word prefixed with *N-* is most commonly interpreted as the actor, but may sometimes be understood as bearing other thematic roles, including that of patient. Thus, in Jakarta Indonesian, the prefixes *di-* and *N-* lack one of the most essential features of prototypical voice markers, namely the function of thematic role assignment.[3]

Thus, as argued above, in Jakarta Indonesian, the prefixes *di-* and *N-* differ in numerous substantive ways from prototypical markers of passive and active voice. So what, then, is the function of these two prefixes? Although not markers of passive and active voice, they are still clearly oriented towards the thematic roles of patient and actor respectively. But rather than having a grammatical function, their function is a straightforward semantic one, that of denoting the relevant thematic role. Thus, the prefix *di-* marks the word to which it attaches, quite simply, as *having* a patient in its semantic argument structure, while the prefix *N-* marks its host word as having an actor in its semantic argument structure.

Constructions containing *di-* and *N-* are of two quite distinct types, depending on the semantic argument structure of the word to which the prefixes attach. Most commonly, the prefixes apply to words denoting activities or properties, whose semantic argument structure already contains the relevant thematic role; in such instances, the effect of the prefix is to *emphasize* the thematic role, adding conceptual saliency. In such instances, the prefixes are, narrowly speaking, redundant, since they do not change the interpretation of the construction per se, but only shade its perspective in a somewhat different way, drawing attention to the presence of the relevant argument.[4]

In this case, the two prefixes are comparable to any number of other optional elements which may be added to a given expression, elaborating on its meaning without any affect on its external grammatical behaviour. For example, like in other so-called "pro-drop" languages, a single word such as *makan* 'eat' may function as a complete sentence; however, an optional pronoun *dia* may be added to emphasize a third person reference, for example *Dia makan* 'He is eating'. Similarly, like in most or all other languages, a single word may be used to refer to an object, such as *mangga* 'mango'; however, an optional word *hijau* 'green' may be added in order to attribute a certain property to the referent, for example *mangga hijau* 'green mango'.

In other instances, however, the prefixes *di-* and *N-* apply to words whose semantic argument structure does not contain the relevant thematic role; in such cases their function is to *introduce* the thematic role into the semantic argument structure of the host word. For example, the word *kunci* 'key' has no participants in its semantic argument structure; adding the prefix *di-* introduces a patient into its semantic argument structure, and, in doing so, changes the meaning of the host word from a thing into an activity 'lock'. Similarly, *kopi* 'coffee' has no participants in its semantic argument structure; adding the prefix *N-* introduces an actor into its semantic argument structure, and changes the meaning of the host word from a thing into an activity 'drink coffee'.

Thus, as argued in this section, the prefixes *di-* and *N-* differ in many substantial ways from prototypical markers of passive and active voice, while still sharing some basic commonalities with such markers, by dint of their orientation towards particular thematic roles, patient and actor respectively. The relationship between the prefixes *di-* and *N-* and prototypical voice markers is most perspicuously understood with reference to a somewhat more generalized notion of voice marker, as laid out in the following definition, first provided in Gil (2002:276-277):

(1) A *generalized voice marker* is a marker M which, when applied to a form X, marks the argument of X bearing the thematic role T as having a set of properties P.

In the above definition, the set of properties P assigned by M to an argument of X is unspecified. This makes it possible to distinguish between different kinds of generalized voice markers, in accordance with the composition of P. In particular, one may distinguish between *strong* generalized voice markers, in which P contains a large number of important properties, and *weak* generalized voice markers, in which P is minimal in size and significance.

Strong generalized voice markers include prototypical markers of passive and active voice. Consider, for example, the English passive marker *be* + *-en*. When applied to its host word, *be* + *-en* marks the argument of its host bearing the thematic role of patient as having a set P of grammatical properties characteristic of subject NPs in English: nominative case (for pronouns), control of verbal agreement, pre-verbal position, and many others. Similarly, the English active marker, which has zero form, marks the argument of its host bearing the thematic role of actor as having the same set P of subject properties.

In contrast, the prefixes *di-* and *N-* provide examples of weak generalized voice markers, in which the set P is minimal in size, comprising just the single basic property of existence. Thus, as argued above, the prefix *di-* marks the argument of its host bearing the thematic role of patient simply as existing: it does not associate it with any further semantic or morphosyntactic properties. Similarly, the prefix *N-* marks the argument of its host bearing the thematic role of actor as existing, without associating it with any additional properties.

Thus, the notion of generalized voice marker provides a unified conceptual framework highlighting both similarities and differences between the prefixes *di-* and *N-* and prototypical markers of passive and active voice. Like prototypical passive and active markers, the prefixes *di-* and *N-* are oriented towards specific thematic roles. However, whereas prototypical passive and active markers have a complex grammatical function, associating

the relevant argument with an array of morphosyntactic and semantic properties, the prefixes *di-* and *N-* possess a simple semantic function, that of asserting the existence of the relevant argument.

To this point, the discussion has focussed on properties shared by the two prefixes *di-* and *N-*. However, these two prefixes also differ from each other in significant ways. Moreover, these differences also set these two prefixes apart from prototypical markers of passive and active voice. In general, across languages, passive forms are more formally complex than their active counterparts. In contrast, the patient-oriented prefix *di-* is actually less formally complex than its actor-oriented counterpart *N-*.

This formal difference between the two prefixes manifests itself in several ways. First, whereas the form of the prefix *di-* is immutable, that of the prefix *N-* is variable, depending on the phonological properties of the first segment of the host word. Secondly, whereas the prefix *di-* is clearly separable from its host word, the prefix *N-* often fuses with the stem through processes of consonant mutation and replacement.

Thirdly, the prefix *di-* is more productive than its counterpart *N-*. Whereas *di-* can attach freely to any word whatsoever, as long as it is semantically appropriate, the distribution of *N-* exhibits a more complex pattern. One constraint is phonological in nature: *N-* occurs considerably less readily before stems beginning with a voiced obstruent: *b*, *d*, *g* or *j*. Beyond that, however, the occurrence of *N-* is lexically conditioned in what can only be described as quite arbitrary fashion. Diachronically, the prefix *N-* may be viewed as entering into usage, establishing a foothold, becoming more and more frequent, until finally it is used invariably, at which point it is reanalyzed as part of the stem, and thus can be said to have disappeared. At any given point in space and time — such as, in the case at hand, present-day Jakarta Indonesian — different words reflect different stages in the ongoing cyclical process of lexicalization; accordingly, the synchronic distribution of *N-* is lacking in semantic motivation. For example, among foodstuffs, whereas *kambing* 'goat' never takes *N-*, *kopi* 'coffee' may. Similarly, among words of visual perception, whereas *lihat* 'see' takes *N-* relatively infrequently, *tonton* 'watch' is already in the process of undergoing reanalysis as *nonton.*

These asymmetries between the prefixes *di-* and *N-* are manifestations of a deeper typological property which Jakarta Indonesian shares with many other related Austronesian languages, that of *patient prominence* (see Cena 1977, De Guzman 1976, 1992, and Gil 1983, 1984). Whereas in most of the world's languages, actors are more important than patients in myriad ways, in patient-prominent Austronesian languages a mirror-image state of affairs typically obtains, whereby patients generally take precedence over actors, in several respects. First, patients tend to be more referentially salient than

actors, for example they are more likely to be definite, more likely to independently refer, more likely to have wide scope, and so on. Secondly, patients tend to be more likely than actors to occur in various specific constructions, for example as the so-called 'topic', 'subject', or 'pivot' of the clause. Finally, patient-oriented generalized voice markers tend to predominate over actor-oriented ones, by exhibiting the following properties: (a) greater formal simplicity; (b) greater productivity; (c) greater ease of processing; (d) greater frequency of occurrence; and (e) earlier acquisition by children. As shown above, the Jakarta Indonesian voice prefixes uphold the first two of these properties, with *di-* exhibiting greater formal simplicity and greater productivity than its counterpart *N-*. And as we shall see in the next section, they also uphold the latter two properties, with *di-* occurring more frequently, and being acquired earlier, than its counterpart *N-*.

3. THE ACQUISITION OF VOICE MORPHOLOGY

In recent years, the acquisition of various dialects of Malay / Indonesian has been the focus of a number of studies. For Jakarta Indonesian, the major previous study is that by Soenjono (1997, 2000) of his own granddaughter.

3.1. The MPI Jakarta Corpus

Since 1999, the Max Planck Institute for Evolutionary Anthropology, in collaboration with Atma Jaya University, have been conducting a large-scale investigation of the acquisition of Jakarta Indonesian. The project involves the longitudinal study of 8 children growing up in Jakarta and acquiring Jakarta Indonesian, and is centered around the compilation of an extensive corpus of naturalistic data. Some previous studies of the acquisition of Jakarta Indonesian based on data from the MPI Jakarta corpus include Cole, Gil, Hermon and Tadmor (2001, to appear), Gil (2003), Hermon and Tjung (2002) and Tadmor (2000, to appear).

The present paper is based on data from the MPI Jakarta corpus. Since the project is ongoing, the present paper makes use of an interim corpus available at the time of writing. A quantitative profile of the interim corpus is provided in Table 1.

Table 1. A Quantitative Profile of the Interim MPI Jakarta Corpus

Target Child		*Age at First Coded Utter.*	*Age at Last Coded Utter.*	*Number of Coded Utter.*
Timo	(M)	1;06	4;00	19,912
Hizkia	(M)	1;07	4;10	24,424
Rizka	(F)	1;08	5;00	32,752
Michael	(M)	2;00	3;07	19,404
Priska	(F)	2;07	4;08	31,642
Larissa	(F)	2;10	5;04	22,280
Ido	(M)	3;04	6;05	31,800
Pipit	(F)	4;04	7;09	25,330
TOTAL: Target children				207,544
Other children				73,477
TOTAL: all children				281,021
TOTAL: all adults				284,062
TOTAL				565,083

As evident from Table 1, the interim corpus contains over 500,000 utterances, of which about one half are from children, comprising the 8 target children plus additional children interacting with them. The remaining half of the utterances are from adults, representing the input to the children's language acquisition.

3.2. The Acquisition of Di- and N-: An Overview

A general overview of the distribution of the two voice prefixes *di-* and *N-* in the interim corpus is provided in Table 2. In Table 2, rows represent age brackets. The first age bracket, 0-1;5, contains relatively little data, coming entirely from children other than the 8 target children. The bulk of the acquisitional data are to be found in the subsequent 9 rows, which break the data down into 6-month intervals, from 1;6-1;11 to 5;6-5;11. The next row provides data for older children, aged 6;0-11;11, the row after that for adults, for present purposes defined as 12;0 or above, while the bottom row adds up all the age brackets for the grand totals. The first two columns specify the number of utterances and words respectively, in each age bracket. The next two columns present the number of occurrences of *di-* and *N-* respectively. The next two columns provide the frequency of *di-* and *N-* respectively, calculated as a prefix-to-word ratio, and represented in percentages. And the final column depicts the relative frequencies of *di-* and *N-*, calculated as a ratio.[5]

Table 2. Occurrence of di- and N- in the Entire Corpus

Age	*Utter.*	*Words*	*di-*	*N-*	*di-/words*	*N-/words*	*di-/N-*
0-1;5	880	1058	0	0	0	0	NA
1;6-1;11	8877	11661	30	11	0.26	0.09	2.73
2;0-2;5	20864	31966	184	104	0.58	0.33	1.77
2;6-2;11	38158	79074	751	334	0.95	0.42	2.25
3;0-3;5	36056	84910	845	453	1.00	0.53	1.87
3;6-3;11	38368	97227	1088	773	1.12	0.80	1.41
4;0-4;5	29923	77853	1004	710	1.29	0.91	1.41
4;6-4;11	28945	81211	1235	771	1.52	0.95	1.60
5;0-5;5	16347	44393	605	348	1.36	0.78	1.74
5;6-5;11	13607	37606	623	272	1.66	0.72	2.29
6;0-11;11	48996	133119	1742	1309	1.31	0.98	1.33
12;0-	284062	802193	14443	8629	1.80	1.08	1.67
Total	565083	1412271	22550	13714	1.60	0.97	1.64

The figures in Table 2 point clearly towards the following three conclusions:

(2) (a) Both *di-* and *N-* are present from an early age.

(b) Both *di-* and *N-* show steady increase in frequency throughout childhood, until adult frequency is obtained.

(c) At all ages, *di-* is more frequent than *N-*.

The first and perhaps most striking conclusion is that the two voicing prefixes, *di-* and *N-*, are present in the corpus from close to the beginning. Thus, in the 1;6-1;11 age bracket, 0.26% of the words are marked with *di-* and 0.09% with *N-*. In fact, given the small amount of data from younger speakers, in the 0-1;5 range, it is not beyond the realm of possibility that a larger corpus would reveal instances of *di-* and *N-* even among speakers under 1;6.[6]

When examining the acquisition of morphological patterns, it is of crucial importance to determine whether the child is creating the forms in question through the application of productive rules, or whether s/he is merely treating them as unanalyzable wholes, perhaps repeating a complex form uttered by an adult in the preceding discourse. Following are some examples of early use of the prefix *di-* which appear to reflect the application of a productive rule of prefixation:

(3) *Dibuka* *Timo 1;8*
di-open
[Asking mother to open picture poster]
'Open it'

cf. *Buka* *Timo 1;7*
open
[Asking mother to open door]
'Open it'

(4) *Buang aja.* *Dibuang.* *Dibuang.* *Timo 1;11*
throw just **di**-throw **di**-throw
[Playing with the microphone]
'Just throw it away. Throw it away. Throw it away.'

(5) *Ditumpahin* *Rizka 1;8*
di-spill-ep
[Seeing another child dropping crackers from a bag]
'She's dropping them'

cf. *Bedak* *tumpah* *Rizka 2;0*
powder spill
[About some face powder]
'The powder's spilt'

(6) *Pasang.* *Dipasang.* *Rizka 1;11*
fasten **di**-fasten
[Watching experimenter set the camera up on the tripod]
'Setting it up. Setting it up'

Examples (3) and (4) are from Timo. In (3), Timo at 1;8 applies the prefix *di-* to the stem *buka*. In the preceding discourse there is no occurrence of the complex form *dibuka* that he could be suggested to be merely echoing. In fact, at 1;7, Timo has already provided evidence for having acquired the bare form *buka*. Thus, it would appear that at 1;8, Timo can already apply the prefix *di-* to the stem *buka* in an appropriate fashion. In (4), Timo at 1;11 shows that he can produce the bare form *buang* alongside the prefixed form *dibuang*, neither of which occur in the preceding discourse. This accordingly provides all-in-one evidence that he has mastered the rule of prefixation. Examples (5) and (6) are from Rizka. In (5), Rizka at 1;8 applies the prefix *di-* (and also the applicative or "end-point" suffix *-in*) to the stem *tumpah*. Again, in the preceding discourse there is no occurrence of complex forms

such as *ditumpah* or *ditumpahin* that she could be suggested to be mimicking. Unfortunately, the corpus does not contain any prior instances of Rizka using the bare form *tumpah*, however, it is very probable that this is no more than an accidental gap in the data, and that she is indeed already aware that the complex form *ditumpahin* contains the prefix *di-* (and, for that matter, also the suffix *-in*). In (6), Rizka at 1;11 uses the bare form *pasang* and then the prefixed form *dipasang*. In this case, both bare and prefixed forms have been uttered relatively recently, by the experimenter setting the camera up; however, rather than suggesting mimicry, this illustrates precisely the kind of input that might be characterized as a naturally-occurring grammar lesson, which, in (6), Rizka shows that she has already mastered. Thus, examples (3) – (6) suggest that two of the children, at least, have mastered the productive use of the patient-oriented prefix *di-* at a very early age.[7]

The next set of examples illustrates early productive occurrences of the prefix *N-*:

(7) *Mbuka* *Timo 1;9*
N-open
[Asking mother to take his shirt off]
'Take it off '

cf. *Buka* *Timo 1;7*
open
[Asking mother to open door]
'Open it'

(8) *xx ngeliat xx* *Timo 1;11*
xx **N-**see xx
[Asking to see something]
'I want to see'

cf. *liat xx* *Timo 1;11*
see xx
[Asking to see something, again]
'I want to see'

(9) *Apa, lagi ngapain?* *Rizka 2;0*
what prog **N-**what-ep
[About a picture in a book]
'What's he doing?'

(10) *Ngambil buah* *Rizka 2;1*
N-take fruit
[Looking at a picture book with mother; mother points to a giraffe and asks what he's doing]
'Taking fruit'

cf. *Ica ambil* *Rizka 2;1*
Ica take
[Asking for a pen that her mother is holding]
'I want to take it'

Again, (7) and (8) are from Timo. In (7), Timo at 1;9 applies the prefix *N-* to the stem *buka*. The preceding discourse contains no occurrence of *mbuka* that Timo could be said to be imitating; rather, the previous few utterances contain a few instances of *dibuka*, with the *di-* prefix. Moreover, as already shown in (3), Timo is familiar with the bare stem *buka*. Thus, it seems as though at 1;9, Timo can successfully and appropriately apply the prefix *N-* to the stem *buka*. In (8), Timo at 1;11 applies the prefix *N-* to the stem *liat*. In this case, the preceding discourse contains a few instances of the bare stem *liat*, and then, when the experimenter fails to understand the child's utterances and asks him to repeat himself, Timo himself produces the bare form *liat*, thereby proving his mastery of the prefix *N-* (if not his ability to communicate whatever it was that he was trying to say). And once more, (9) and (10) are from Rizka. In (9) Rizka at 2;0 produces side-by-side interrogatives based on *apa*, the first bare, the second prefixed with *N-* (and also suffixed with the applicative *-in*), and, in doing so, demonstrates her understanding of the relationship between the two. Finally, in (10), Rizka at 2;1 applies the prefix *N-* to the stem *ambil*, having, just a couple of weeks before, produced the bare stem form *ambil*. In the case at hand, the preceding discourse contains no occurrences of the prefixed form *ngambil*. Interestingly, though, Rizka's use of it, in (10), is in immediate response to her mother's question using the interrogative *ngapain* containing the prefix *N-*. Often, in adult discourse, the interrogative form *ngapain* 'do what' is answered with an activity word also prefixed with *N-*; thus, Rizka's use of the prefix *N-* in her answer shows that she has mastered this characteristic discourse pattern involving the prefix *N-*. Thus, examples (7) - (10) suggest

that the same two children, at least, have also mastered the productive use of the prefix *N-* at an early age.

Thus, as suggested by Table 2 and examples (3) - (10), both *di-* and *N-* occur in productive use from an early age. Not surprisingly, however, their use starts out as less frequent, increasing as the children grow older. This increase in frequency continues through at least part of the older-child 6;0-11;11 age bracket, until the adult frequencies of 1.80% and 1.08% respectively are achieved.

The early use of the patient-oriented voice marker *di-* in Jakarta Indonesian thus places it in the company of passive markers in Zulu, Sesotho, Tagalog, Inuktitut and Quiche, mentioned in the introduction, while distinguishing it from passives in English and other languages, which are typically acquired at a later stage. How might one account for the early acquisition of *di-* in Jakarta Indonesian? Two explanations suggest themselves, one structural, the other statistical.

On the one hand, one might attribute the early acquisition of *di-* to its formal simplicity. As pointed out in Section 2, passives in English and other similar languages are strong voice markers, with a relatively complex grammatical function, assigning the patient argument an array of morphosyntactic properties characteristic of subjects in the respective languages. In contrast the prefix *di-* is a weak voice marker, associated with a relatively simple semantic function, namely asserting the existence of a patient argument. Thus, children might be acquiring the prefix *di-* earlier because it is simpler; in order to use it appropriately they do not have to figure out any of the grammatical complexities associated with prototypical passives. On the other hand, it would seem just as plausible to account for the early acquisition of *di-* in terms of its greater frequency in the adult input. As indicated in Table 2, 1.80% of all words in the adult input contain the prefix *di-*. In contrast, the corresponding figures for English and many other languages are much lower; see for example Svartvik (1966). Thus, children might be producing the prefix *di-* earlier because they are exposed to it more often. This explanation is consistent with a view of children as being "little statistical machines" that calculate frequencies in the input and make use of the results to modulate their own linguistic behaviour. Thus, simplicity and frequency would seem to be two good reasons for the early acquisition of *di-*, and, indeed, in Section 3.3 below, evidence is provided that the early acquisition of *di-* is due to an interplay of these two factors.

Turning now to the last column of Table 2, an additional interesting fact emerges: from early childhood all the way through to adulthood, the prefix *di-* occurs with greater frequency than its counterpart *N-*.[8] Again, the acquisitional facts may be explained in two different ways, referring respectively to structure and to statistics.

On the one hand, one might attribute the greater frequency of *di-* in the children to the patient prominence of Jakarta Indonesian and concomitant greater formal simplicity of the prefix *di-*. Thus, in order to use the prefix *N-* properly, the child needs to master three additional and more complex tasks: to figure out the appropriate morphophonemic realization of the prefix, to fuse it in the correct way to the stem, and to verify that the stem in question is indeed one that may occur with the prefix *N-*. No wonder, then, that they have a harder time using *N-* than its simpler counterpart *di-*. On the other hand, it would seem just as reasonable to attribute the greater frequency of *di-* in the children to the greater frequency of *di-* in the adult input. As indicated in the last-but-one row of Table 2, among adults *di-* is 1.67 times more frequent than *N-*. Perhaps, then, children are able to directly access the *di- / N-* ratio in adult speech in order to recreate it in their own. Thus, once again, formal simplicity and adult frequency would seem to be two equally cogent reasons why children use *di-* more frequently than *N-*. As we shall now see, both of these two factors are probably responsible, in tandem, for the observed facts.

3.3. The Acquisition of Di- and N- by Individual Children

In order to obtain a more nuanced picture of the acquisition of *di-* and *N-*, it is necessary to examine the process of acquisition as it plays out in individual children. Since these two prefixes enter into usage at a very early age, we shall limit our attention to the four target children for whom the corpus contains the earliest data: Timo, Rizka, Michael and Hizkiah.[9]

Tables 3-6 show the distribution of the prefixes *di-* and *N-* in the speech of Timo, Rizka, Michael and Hizkiah respectively, divided into 3-month intervals. The layout of these tables is identical to that of Table 2. Tables 7-9 extract the data from the last three columns of Tables 3-6 and present it in a format conducive to easy comparisons across children. Thus, Table 7 summarizes the frequency of *di-* for the four children, Table 8 presents the frequency of *N-* for the same children, while Table 9 compares their *di- / N-* frequency ratios.

Table 3: Occurrence of di- and N- in Timo

Age	*Utter.*	*Words*	*di-*	*N-*	*di- /words*	*N- /words*	*di- /N-*
1;6-1;8	1347	1324	4	0	0.30	0	∞
1;9-1;11	820	1111	11	4	0.99	0.36	2.75
2;0-2;2	1820	2624	18	8	0.69	0.30	2.25
2;3-2;5	2316	4067	48	14	1.18	0.34	3.43

2;6-2;8	1989	3945	53	7	1.34	0.18	7.57
2;9-2;11	2820	5396	34	19	0.63	0.35	1.79
3;0-3;2	2560	5485	38	14	0.69	0.26	2.71
3;3-3;5	1444	3436	50	9	1.46	0.26	5.56
3;6-3;8	2533	5556	34	24	0.61	0.43	1.42
3;9-3;11	2022	4511	53	16	1.17	0.35	3.31
4;0-4;2	241	443	4	2	0.90	0.45	2.00
Total	19912	37898	347	117	0.92	0.31	2.97

Table 4. Occurrence of di- and N- in Rizka

Age	*Utter.*	*Words*	*di-*	*N-*	*di- /words*	*N- /words*	*di- /N-*
1;6-1;8	96	195	1	0	0.51	0	∞
1;9-1;11	976	1557	11	2	0.71	0.13	5.50
2;0-2;2	1849	3349	35	16	1.05	0.48	2.19
2;3-2;5	1630	3151	39	14	1.24	0.44	2.79
2;6-2;8	3524	8340	99	81	1.19	0.97	1.22
2;9-2;11	3537	11291	196	86	1.74	0.76	2.28
3;0-3;2	2486	6496	79	49	1.22	0.75	1.61
3;3-3;5	1576	4484	65	37	1.45	0.83	1.76
3;6-3;8	2910	8414	84	108	1.00	1.28	0.78
3;9-3;11	2358	6763	82	85	1.21	1.26	0.96
4;0-4;2	2515	7827	103	80	1.32	1.02	1.29
4;3-4;5	2621	7803	105	85	1.35	1.09	1.24
4;6-4;8	3140	9423	100	114	1.06	1.21	0.88
4;9-4;11	2642	8106	128	104	1.58	1.28	1.23
5;0-5;2	892	2723	34	33	1.25	1.21	1.03
Total	32752	89922	1161	894	1.29	0.99	1.30

Table 5. Occurrence of di- and N- in Michael

Age	*Utter.*	*Words*	*di-*	*N-*	*di- /words*	*N- /words*	*di- /N-*
2;0-2;2	3310	3486	4	13	0.11	0.37	0.31
2;3-2;5	2012	3213	5	19	0.16	0.59	0.26
2;6-2;8	2430	5315	22	18	0.41	0.34	1.22
2;9-2;11	2190	5691	35	9	0.62	0.16	3.89
3;0-3;2	3732	9454	88	28	0.93	0.30	3.14
3;3-3;5	3367	7167	53	21	0.74	0.29	2.52
3;6-3;8	2363	5248	46	28	0.88	0.53	1.64
Total	19404	39574	253	136	0.64	0.34	1.86

Table 6. Occurrence of di- and N- in Hizkiah

Age	*Utter.*	*Words*	*di-*	*N-*	*di-/words*	*N-/words*	*di-/N-*
1;6-1;8	229	279	0	0	0	0	NA
1;9-1;11	1406	2230	0	0	0	0	NA
2;0-2;2	2425	3383	2	1	0.06	0.03	2.00
2;3-2;5	1041	1721	0	0	0	0	NA
2;6-2;8	2663	4562	10	5	0.22	0.11	2.00
2;9-2;11	2390	4846	36	12	0.74	0.25	3.00
3;0-3;2	1582	3648	55	18	1.51	0.49	3.06
3;3-3;5	1696	4267	49	33	1.15	0.77	1.48
3;6-3;8	2537	6550	86	24	1.31	0.37	3.58
3;9-3;11	2132	5570	80	38	1.44	0.68	2.11
4;0-4;2	2119	5677	40	52	0.70	0.92	0.77
4;3-4;5	1716	4948	57	29	1.15	0.59	1.97
4;6-4;8	1523	5020	77	49	1.53	0.98	1.57
4;9-4;11	965	3270	26	10	0.80	0.31	2.60
Total	24424	55971	518	271	0.93	0.48	1.91

Table 7. di- / Word Ratio in the Four Youngest Children

Age	*Timo*	*Rizka*	*Michael*	*Hizkiah*
1;6-1;8	0.30	0.51		0
1;9-1;11	0.99	0.71		0
2;0-2;2	0.69	1.05	0.11	0.06
2;3-2;5	1.18	1.24	0.16	0
2;6-2;8	1.34	1.19	0.41	0.22
2;9-2;11	0.63	1.74	0.62	0.74
3;0-3;2	0.69	1.22	0.93	1.51
3;3-3;5	1.46	1.45	0.74	1.15
3;6-3;8	0.61	1.00	0.88	1.31
3;9-3;11	1.17	1.21		1.44
4;0-4;2	0.90	1.32		0.70
4;3-4;5		1.35		1.15
4;6-4;8		1.06		1.53
4;9-4;11		1.58		0.80
5;0-5;2		1.25		
Total	0.92	1.29	0.64	0.93

Table 8. N- / Word Ratio in the Four Youngest Children

Age	*Timo*	*Rizka*	*Michael*	*Hizkiah*
1;6-1;8	0	0		0
1;9-1;11	0.36	0.13		0
2;0-2;2	0.30	0.48	0.37	0.03
2;3-2;5	0.34	0.44	0.59	0
2;6-2;8	0.18	0.97	0.34	0.11
2;9-2;11	0.35	0.76	0.16	0.25
3;0-3;2	0.26	0.75	0.30	0.49
3;3-3;5	0.26	0.83	0.29	0.77
3;6-3;8	0.43	1.28	0.53	0.37
3;9-3;11	0.35	1.26		0.68
4;0-4;2	0.45	1.02		0.92
4;3-4;5		1.09		0.59
4;6-4;8		1.21		0.98
4;9-4;11		1.28		0.31
5;0-5;2		1.21		
Total	0.31	0.99	0.34	0.48

Table 9: Di- / N- Ratio in the Four Youngest Children

Age	*Timo*	*Rizka*	*Michael*	*Hizkiah*
1;6-1;8	∞	∞		NA
1;9-1;11	2.75	5.50		NA
2;0-2;2	2.25	2.19	0.31	2.00
2;3-2;5	3.43	2.79	0.26	NA
2;6-2;8	7.57	1.22	1.22	2.00
2;9-2;11	1.79	2.28	3.89	3.00
3;0-3;2	2.71	1.61	3.14	3.06
3;3-3;5	5.56	1.76	2.52	1.48
3;6-3;8	1.42	0.78	1.64	3.58
3;9-3;11	3.31	0.96		2.11
4;0-4;2	2.00	1.29		0.77
4;3-4;5		1.24		1.97
4;6-4;8		0.88		1.57
4;9-4;11		1.23		2.60
5;0-5;2		1.03		
Total	2.97	1.30	1.86	1.91

In order to identify the relevant factors underlying the acquisition of the prefixes *di-* and *N-* by the four selected children, it is also necessary to

examine the distribution of these two prefixes differentially in the adult input for each individual child. Such a profile is provided in Table 10.

Table 10. Occurrence of di- and N- in the Adult Input to the Four Youngest Children

Adult input	*Timo*	*Rizka*	*Michael*	*Hizkiah*
Utterances	40311	47363	28795	31807
Words	98524	141193	81067	89424
di-	1570	2770	1212	1695
N-	741	1983	702	888
di- / words	1.59	1.96	1.50	1.90
N- / words	0.75	1.40	0.87	0.99
di- / N-	2.12	1.40	1.73	1.91

Examination of the figures in Tables 3-9 reveals significant differences between the children, with respect to their rates of development. These differences can be seen most clearly in Tables 7 and 8. To aid in the description of these differences, we may posit four stages in the acquisition of each of the two prefixes:[10]

(11) (a) Stage 1:
Voice marker absent;

(b) Stage 2:
Voice marker increasing rapidly in frequency;

(c) Stage 3:
Voice marker increasing gradually in frequency;

(d) Stage 4:
Voice marker stable at adult frequency.

The first observation to be made is that none of the four children achieve Stage 4 in the course of the study. This can be seen by comparing the later figures for each child in Tables 7 and 8 with the adult figures from Table 2, 1.80% for *di-* and 1.08% for *N-*.[11] The second observation is that in some cases, the study started too late to capture Stage 1, in which the prefixes are absent. Thus, for *di-*, only Hizkiah is seen in Stage 1, while for *N-*, only Timo, Rizka and Hizkiah are observed in Stage 1.

Of interest is Stage 2: its onset from Stage 1, and its transition to Stage 3. To assist in the visual processing of the figures in Tables 7 and 8, wavy lines are placed at the approximate boundaries between Stages 1 and 2, and between Stages 2 and 3.[12] Thus for each child-prefix pair, or column in the table, two wavy lines segment the figures into three chunks, reflecting Stages 1, 2 and 3

respectively. The results of this segmentation are summarized in Tables 11 and 12 (with "..." indicating presumed additional age brackets outside the scope of the corpus).

Table 11: Stages in the Acquisition of di- in the Four Youngest Children

Child	*Timo*	*Rizka*	*Michael*	*Hizkiah*
Stage 1				… 2;5
Stage 2	… 1;6-1;8	… 1;6-1;11	… 2;0-2;8	2;6-2;8
Stage 3	1;9-4;2 …	2;0-5;2 …	2;9-3;8 …	2;9-4;11 …

Table 12: Stages in the Acquisition of N- in the Four Youngest Children

Child	*Timo*	*Rizka*	*Michael*	*Hizkiah*
Stage 1	… 1;8	… 1;8		… 2;5
Stage 2	short	1;9-2;5		2;6-2;11
Stage 3	1;9-4;2 …	2;6-5;2 …	2;0-3;8 …	3;0-4;11 …

Thus, Tables 11 and 12 show, for *di-* and *N-* respectively, the development of each child, by tracing the three stages of acquisition with respect to the 3-month age intervals.

The relative precociousness of the four children, as evidenced in Tables 11 and 12, may be summarized as follows:

(12) Relative precociousness of the four youngest children
  (a) for *di-* Timo > Rizka > Michael > Hizkiah
  (b) for *N-* [Michael] Timo > Rizka > Hizkiah

In Table 11, for *di-*, the four children are arranged in order of decreasing precociousness from left to right, as recapitulated in (12a). Thus, onset of Stage 2 is sometime before 1;6 for Timo and Rizka, sometime before 2;0 for Michael, but 2;6 for Hizkiah. Similarly, onset of Stage 3 is 1;9 for Timo, 2;0 for Rizka, and 2;9 for Michael and Hizkiah. A similar though not identical picture emerges in Table 12, for *N-*, as shown in (12b); here, Timo, Rizka and Hizkiah again exhibit a pattern of decreasing precociousness, however, Michael moves up somewhat in the hierarchy, though his precise position is impossible to determine, due to absence of data for the earlier stages. Thus, onset of Stage 2 is 1;9 for Timo and Rizka, sometime before 2;0 for Michael, and 2;6 for Hizkiah. And onset of Stage 3 is 1;9 for Timo, sometime before 2;0 for Michael, 2;6 for Rizka, and 3;0 for Hizkiah.

The similarities between the patterns for *di-* and *N-* suggest that the differences between the children are not random but rather systematic. Where might such differences come from? One obvious hypothesis is that the earlier

acquisition of the prefixes might be the product of greater frequency of occurrence in the adult input. However, the relevant data, in Table 10, clearly refute this hypothesis. As shown in Table 10, the frequencies of the two prefixes across the respective inputs to the four children point towards the following rankings:

(13) Frequencies in the adult input of the four youngest children
  (a) of *di-* Rizka ≥ Hizkiah > Timo ≥ Michael
  (b) of *N-* Rizka > Hizkiah > Michael > Timo

Comparing (13a) to (13b) shows that the two prefixes exhibit similar rankings with respect to their frequency in the adult input, the only substantive difference being that of the relative positions of Timo and Michael. However, comparing (13) to (12) demonstrates that there is no correlation between frequency in the adult input and earliness of acquisition. For example, Timo exhibits early acquisition of both prefixes in spite of their relatively low frequencies in the adult input, 1.59% and 0.75% respectively, whereas Hizkiah is the slowest to acquire both prefixes in spite of a relatively larger exposure to them in the adult input, 1.90% and 0.99% respectively. Thus, the different rates of acquisition of the two prefixes by the four children, as summarized in (12) above, cannot be due to differences in the frequencies of these prefixes in their respective adult inputs. At this point, one can only speculate as to what the factors are that underlie the variation between children in the initial acquisition of these two prefixes.[13]

Nevertheless, one important conclusion can be drawn from the above discussion. If, as is the case here, the rate of acquisition of the two prefixes is at least partly independent of input frequency across children within the same language, then it stands to reason that the rate of acquisition of voice markers should be at least partly independent of input frequency across different languages as well. Recall that in the previous subsection it was suggested that the earlier acquisition of the prefix *di-* compared to passive markers in English and other languages might be due either to its formal simplicity or to its frequent occurrence. The above discussion accordingly suggests that the high frequency of *di-* relative to passive markers in English and other languages cannot be the sole predicator of its earlier acquisition. By elimination, it thus points to the conclusion that the early acquisition of *di-* must be due, at least in part, to its structural properties, namely its formal simplicity.

However, if we consider a somewhat different aspect of the four children's acquisition of the two prefixes, then it turns out that frequency of input does indeed play a role. Table 13 compares the occurrence of the two prefixes in the four children's speech during an identical one-year time span,

2;9-3;8, almost entirely contained within Stage 3 of development (the only exception being Hizkiah who at 2;9-2;11 is still in Stage 2 for *N*-).

Table 13: Occurrence of di- and N- in the Four Youngest Children at Age 2;9-3;8

Child	*Timo*	*Rizka*	*Michael*	*Hizkiah*
Utterances	9357	10509	11652	8205
Words	19873	30685	27561	19311
di-	156	424	222	226
N-	66	280	86	85
di- / words	0.78	1.38	0.81	1.17
N- / words	0.33	0.91	0.31	0.44
di- / N-	2.36	1.51	2.58	2.66

Table 13 suggests the following ranking of the four children with respect to the frequency of occurrence of the two prefixes in the specified age bracket:

(14) Frequencies in the four youngest children at age 2;9-3;8
  (a) of *di-* Rizka > Hizkiah > Michael ≥ Timo
  (b) of *N-* Rizka > Hizkiah > Timo ≥ Michael

Comparing (14a) to (14b) shows that the two prefixes exhibit similar rankings with respect to their frequency in the four children, the only slight difference being that of the relative positions of Timo and Michael. However, comparing (14) to (12) demonstrates that there is no correlation between frequency in the specified age bracket and earliness of acquisition. For example, Timo who starts off early with both prefixes, at 2;9-3;8 has only reached a relatively low 0.78% and 0.33%, while Hizkiah, the latest starter, has already soared ahead with 1.17% and 0.44% for *di-* and *N-* respectively. On the other hand, comparing (14) to (13) reveals a near perfect match. What this means is that by age 2;9-3;8, or, perhaps more perspicuously put, by some point early in Stage 3 of development, the differential frequencies of the two prefixes in the children's speech mirror those present in the respective adult inputs, even though the actual frequencies of the prefixes are still much lower in the children than in the adults. (If the latter were not the case, they would, by definition, already be in Stage 4.) For example, Hizkiah's high 1.17% and 0.44% match the high figures from Hizkiah's input, 1.90% and 0.99%, while Timo's low 0.78% and 0.33% reflect the low figures from Timo's input, 1.59% and 0.75%, for *di-* and *N-* respectively. Thus, even though the age of initial acquisition of the two prefixes is determined by

other factors, once the prefixes kick in, frequencies in the adult input do play a role, after all, in the ongoing developmental processes.

Once more, patterns of variation across children bear implications with regard to patterns of variation across languages. If, just a few paragraphs back, it was argued that the early acquisition of *di-* in comparison to markers of passive in English and other languages must be due, at least in part, to structural properties of the prefix, the above discussion suggests that relative frequencies may also play a role, and that the early acquisition of *di-* in Jakarta Indonesian might also be due to its higher frequency in the adult language, in comparison to passive markers in English and other languages. Thus, the emerging picture is one in which formal simplicity and frequency of occurrence both play a significant role in the acquisition of voice morphology, each contributing to the patterns of cross-linguistic variation that are in evidence.

4. CONCLUSION

As shown in this paper, children acquire the two voice prefixes *di-* and *N-* at a very early age, with some children observed using them productively and appropriately under the age of 2;0. The frequency of the two prefixes increases steadily throughout childhood until adult frequency is obtained. At all ages, the patient-oriented prefix *di-* is more frequent than its actor-oriented counterpart *N-*.

The early acquisition of the two voice prefixes is accounted for in terms of two distinct factors. First, these prefixes are formally simple, with a straightforward semantic function devoid of any grammatical complexity. Secondly, these prefixes occur with high frequency in the adult input. Variable patterns of acquisition across the different children makes it possible to tease these two factors apart and demonstrate that each of them separately plays an important role in the acquisition of the two prefixes.

This study adds to an increasing body of research documenting the existence of cross-linguistic variation in the acquisition of voice and passive morphology, showing that in many languages, passive constructions are acquired earlier than in English. Although the data from these various studies is not directly comparable, it would seem to be the case that Jakarta Indonesian children are even more precocious in their use of the passive than children acquiring some of the other languages, such as Sesotho, Inuktitut and Quiche.

For Inuktitut, Allen and Crago (1993, 1996) speculate that the early acquisition of passives may be due either to their structural simplicity or to their high frequency in the adult input. The present study suggests that Allen and Crago's speculations are well-founded, and that both factors are indeed

relevant. The results of this paper thus support an eclectic approach in which diverse and heterogeneous factors are acknowledged to play a role in the acquisition of language.

Finally, the early acquisition of passive constructions has implications with respect to the maturation hypothesis proposed by Borer and Wexler (1987), which suggests that "the process of A-chain formation underlying NP-movement", a necessary component of the passive construction, matures at around 4;0. Allen and Crago (1993, 1996) argue that the early passives of Inuktitut and other languages cast doubt on the maturation hypothesis, since they assume that in Inuktitut, the passive construction involves NP movement. But whatever the most appropriate analysis of the passive construction might be for Inuktitut or for that matter English, there is no positive evidence whatsoever that Jakarta Indonesian has such a rule of NP movement, in constructions involving the patient-oriented prefix *di-* or anywhere else. Suffice it to recall that in Jakarta Indonesian, the presence of the voice prefixes *di-* and *N-* has no effect on word order. Thus, the early acquisition of voice morphology in Jakarta Indonesian as documented in this paper has no bearing whatsoever on the maturation hypothesis.

DAVID GIL *MPI for evolutionary Anthropology, Leipzig*

5. ACKNOWLEDGMENTS

This paper would not have been possible without the unflagging and good-spirited cooperation of Hizkiah, Ido, Larissa, Michael, Pipit, Priska, Rizka and Timo, and their parents, families and friends. It would not have been possible without years of dedicated work by the MPI Jakarta Field Station staff: coordinator Uri Tadmor; information systems manager Brad Taylor; data editors Sarah Chakrawati, Liana Husain, Betty Litamahuputty, Guy Sharett and Antonia Soriente; and research assistants Dalan Mehuli, Dini Andarini, Erni Farida Sri Ulina Ginting, Ferdinand Okki Kurniawan, Lanny Hidayat, Like Raskova Octaberlina, Regina Yanti, Widya Setiabudi, and Yokebed Triwigati. It would not have been possible without the collaboration of our host institution Atma Jaya University and the staff at the Pusat Kajian Bahasa dan Budaya, and in particular Bambang Kaswanti Purwo and Soenjono Dardjowidjojo. And it would not have been possible without the support of Bernard Comrie. Thank you all! This paper has benefited from helpful comments and suggestions by Sharon Armon-Lotem, Peter Cole, Gabriella Hermon, Uri Tadmor, Yassir Tjung, and participants at the Conference on Acquisition of Verb Grammar and Verb Arguments, Zentrum für Algemeine Sprachwissenschaft, Berlin, Germany, 16-17 November 2001. This paper is an abridged version of Gil (to appear b).

6. NOTES

[1] Similar word order flexibility in the Riau dialect of Indonesian is exemplified and discussed in more detail in Gil (1994:181; 1999:191-193; 2002:246-248).

[2] Examples of constructions containing such doubly-marked forms are provided in Gil (to appear b). In addition, similar examples from Riau Indonesian are discussed in Gil (1999:205-206; 2002:265). In discussions of examples such as these, it is sometimes suggested that the prefix *N-* has undergone reanalysis as part of the stem, and hence, that such constructions actually involve only a single instance of prefixation, with *di-*. As noted towards the end of this subsection, the prefix *N-* may indeed undergo reanalysis, but this is the case only for an arbitrary and restricted subset of lexical items. Moreover, the stem *tempel* occurs frequently in bare form in Jakarta Indonesian. Thus, in *dinempel*, the prefix *N-* has not undergone reanalysis as part of the stem; hence, *dinempel* does in fact constitute a bona fide instance of productive double prefixation, with both *di-* and *N-*.

[3] Examples of constructions in which NPs preceding words prefixed with *di-* and *N-* are assigned "unexpected" thematic roles are provided in Gil (2002:268-269; to appear b). In addition, similar examples from Riau Indonesian are discussed in Gil (1999:193-195,200-201; 2002:249-250,259-261).

[4] More detailed discussion and illustration of how these prefixes shade the perspective of such constructions in a similar way in the Riau dialect of Indonesian is provided in Gil (1999, 2002).

[5] The representation of the frequency of *di-* and *N-* in terms of a form-to-word ratio, rather than, as is more customary, a form-to-clause ratio, is for a combination of principled, methodological and practical reasons. On principled grounds, it is not clear whether the syntactic organization of Jakarta Indonesian makes reference to a notion of clause that is comparable to that familiar from many other languages. On methodological grounds, I am not familiar with any rule of thumb that would enable such units, whatever their theoretical validity, to be clearly, systematically and consistently identified. Finally, on practical grounds, even if a workable algorithm for identifying clauses were available, its application to a corpus of this size would have involved either a major undertaking in automatic tagging or a huge effort in manual coding. Alternatively, it would have been easy to represent the frequency of the two prefixes in terms of a form-per-hour rate, as is occasionally done in other studies, for example Allen and Crago (1993, 1996). However, it is unclear how such a calculation would control for such obviously extraneous factors as how talkative the subject children happen to be.

[6] These results are consistent with those of Soenjono (1997:4-5, 2000:121), who found that his granddaughter used the prefix *di-* productively at the age of 1;9.

[7] Further reason to believe that children producing forms containing *di-* are generally aware of their internal structure derives from phonotactic considerations. In Jakarta Indonesian, the canonical and most frequent word structure is bisyllabic. (In an analysis of word structure in the related Riau dialect in Gil to appear a, it is claimed that word structure is centered around a core bisyllabic foot; many of the arguments presented there carry over also to the Jakarta dialect.) Presumably, children acquire this basic rhythmic property of words at a very early age. Observe, now, that *di-* prefixation generally produces trisyllabic words. Such atypically long words should thus strike the language learner as unusual, and provide further motivation for the child to seek an analysis decomposing them into a canonical bisyllabic stem plus a monosyllabic prefix.

[8] The numerical predominance of *di-* over *N-* is probably greater than that suggested by the figures in the last row of Table 2, since, as pointed out in Section 2, the prefix *N-* is prone to lexicalization, leading ultimately to its reanalysis as part of the stem. Thus, the figures for *N-* in Table 2 (and subsequent tables) include forms such as *nonton*, *nangis* and *ngantuk*, where the prefix *N-* is on the verge of reanalysis.

[9] Whereas in Table 1, the children are listed in order of age at the beginning of their respective longitudinal studies, from this point on the four youngest children are presented in an

order which reflects their relative precociousness in the acquisition of the prefix *di-*, as represented in (12a) below.

[10] Whereas the transition from Stage 1 to Stage 2 is abrupt, those from Stage 2 to Stage 3 and from Stage 3 to Stage 4 are gradual and potentially protracted. These four stages may thus be viewed as discrete idealizations of a mostly continuous path of development.

[11] A possible exception to this claim is provided by Rizka in her usage of the prefix *N-*. Although she has already reached the overall adult frequency of *N-*, 1.08%, the higher frequency of *N-* in her own adult input, 1.40%, make it plausible that she is heading for a yet higher frequency than that which she has achieved in the present corpus.

[12] In Table 7, for Timo, Rizka, and Michael, the first wavy line is placed above the first filled cell, indicating that the onset of Stage 2 occurs at some point before data is available. In Table 8, for Timo, the double wavy line between 1;6-1;8 and 1;9-1;11 represents a very rapid transition from Stage 1 through Stage 2 to Stage 3, while for Michael, the double wavy line above the first filled cell suggests that he has already entered Stage 2 and then Stage 3 before data is available.

[13] Some of the relevant factors might include gender, socio-economic status and ethnicity. However, in order to test hypotheses concerning the possible effects of such factors, it would be necessary to have data from a much larger number of children than were available for the present study.

7. REFERENCES

Allen, Shanley E.M. and Martha B. Crago. "Early Acquisition of Passive Morphology in Inuktitut." In Eve Clark (ed)., *Proceedings of the Twenty-Fifth Annual Child Language Research Forum*, Stanford, 1993: 112.

Allen, Shanley E.M. and Martha B. Crago. "Early Passive Acquisition in Inuktitut", *Journal of Child Language* 23 (1996): 129-155.

Baldie, Brian. "The Acquisition of the Passive Voice", *Journal of Child Language,* 3 (1976): 331-348.

Borer, Hagit and Kenneth Wexler. "The Maturation of Syntax", in T. Roeper and E. Williams (eds.), *Parameter Setting*, Reidel, Dordrecht, 1987: 123-172.

Bridges, A. "SVD Comprehension Strategies Reconsidered: The Evidence of Individual Patterns of Response", *Journal of Child Language* 7 (1980): 89-104.

Brown, Roger. *A First Language: The Early Stages*, Cambridge: Harvard University Press, 1973.

Cena, Resty M. "Patient Primacy in Tagalog", *Paper presented at LSA Annual Meeting*, Chicago, 28-30 December, 1977.

Cole, Peter, David Gil, Gabriella Hermon and Uri Tadmor. "The Acquisition of In-Situ WH-Questions and WH-Indefinites in Jakarta Indonesian", in A. H.-J. Do, L. Domínguez and A. Johansen (eds.), *Proceedings of the 25th Annual Boston University Conference on Language Development*, Volume 1, Somerville: Cascadilla Press, 2001: 169-179.

Cole, Peter, David Gil, Gabriella Hermon and Uri Tadmor. "The Acquisition of WH Forms in Jakarta Indonesian, A Preliminary Study", in Uri Tadmor (ed.), *Studies in the Acquisition of Malay / Indonesian, NUSA Linguistic Studies in Indonesian and Languages in Indonesia*, Jakarta: Universitas Atma Jaya, (to appear).

De Guzman, Videa P. *Syntactic Derivation of Tagalog Verbs*, PhD Dissertation, Honolulu: University of Hawaii, 1976.

De Guzman, Videa P. "Morphological Evidence for the Primacy of Patient as Subject in Tagalog", in M. Ross (ed.), *Papers in Austronesian Linguistics*, No. 2. Pacific Linguistics Series A -82, Canberra: Australian National University, 1992: 87-96.

Demuth, Katherine. "Maturation and the Acquisition of Sesotho Passive", *Language* 65 (1989): 56-80.

Demuth, Katherine. "Subject, Topic and Sesotho Passive", *Journal of Child Language* 17 (1990):67-84.

Gil, David. "Quantifier Scope and Patient Prominence: Some Data from Batak, Buginese, and Tagalog", in S. Mordechay (ed.), *UCLA Working Papers in Semantic*, Los Angeles, 1983: 315-364.

Gil, David. "On the Notion of 'Direct Object' in Patient Prominent Languages", in Franz Plank (ed.), *Objects: Towards a Theory of Grammatical Relations*, London: Academic Press, 1984: 87-108.

Gil, David. "The Structure of Riau Indonesian", *Nordic Journal of Linguistics* 17 (1994): 179-200.

Gil, David. "Riau Indonesian as a Pivotless Language", in Elena V. Raxilina and Yakov G. Testelec (eds.), *Tipologija i Teorija Jazyka, Ot Opisanija k Objasneniju, K 60-Letiju Aleksandra Evgen'evicha Kibrika (Typology and Linguistic Theory, From Description to Explanation, Forthe 60th Birthday of Aleksandr E. Kibrik),* Moscow: Jazyki Russkoj Kul'tury, 1999: 187-211.

Gil, David. "The Prefixes di- and N- in Malay / Indonesian Dialects", in F. Wouk and M. Ross (eds.), *The History and Typology of Western Austronesian Voice Systems*, Canberra: Pacific Linguistics, 2002: 241-283.

Gil, David. "How Children Acquire Syntactic Categories in Jakarta Indonesian", in Katharina Endriati Sukatmo (ed.), *Kolita 1, Konferensi Linguistik Tahunan Atma Jaya, Pusat Kajian Bahasa dan Budaya*, Jakarta: Unika Atma Jaya, 2003: 265–269.

Gil, David. "Intonation and Thematic Roles in Riau Indonesian", in M. Gordon, D. Büring and C.M. Lee eds., *Topic, Focus and Intonation [tentative title]*, Deventer: Kluwer, (to appear a).

Gil, David. "The Acquisition of the Prefixes di- and N- in Jakarta Indonesian", in U. Tadmor ed., *Studies in the Acquisition of Malay / Indonesian, NUSA, Linguistic Studies of Indonesian and Other Languages in Indonesia*, Jakarta: Universitas Katolik Indonesia Atma Jaya, (to appear b).

Harwood, F. "Quantitative Study of the Speech of Australian Children"*, Language and Speech* 2 (1959): 237-271.

Hermon, Gabriella and Yassir Tjung. "The Acquisition of Relative Clauses in Colloquial Jakarta Indonesian", Paper presented at *the Sixth International Symposium on Malay / Indonesian Linguistics, Association for Linguistic Typology,* Bintan Island, Indonesia, 4 August 2002.

Horgan, Dianne. "The Development of the Full Passive", *Journal of Child Language* 5 (1978): 65-80.

Mintz, Malcolm W. *A Student's Grammar of Malay and Indonesian*, Singapore: EPB Publishers, 1994.

Pye, Clifton and Pedro Quixtan Poz, "Precocious Passives and Antipassives in Quiche Mayan", *Papers and Reports on Child Language Development* 27 (1988): 71-80.

Segalowitz, Norman S. and Rosita G. Galang. "Agent-Patient Word Order Preference in the Acquisition of Tagalog", *Journal of Child Language* 5 (1978): 47-64.

Soenjono Dardjowidjojo. "The Acquisition of Indonesian: The First Two Years of Life", Paper presented at *the 16th International Congress of Linguists*, Paris, France, 20-25 July 1997.

Soenjono Dardjowidjojo. *Echa, Kisah Pemerolehan Bahasa Anak Indonesia*, Jakarta: Grasindo, 2000.

Suzman, S.M. "Learning the Passive in Zulu", *Papers and Reports on Child Language Development* 24 (1985): 131-137.

Suzman, S.M. "Passives and Prototypes in Zulu Children's Speech", *African Studies* 46 (1987): 241–254.

Svartvik, J. *On Voice in the English Verb*, The Hague: Mouton, 1966.

Tadmor, Uri. "The Acquisition of First and Second Person Pronouns in Jakarta Indonesian", Paper presented at *the 4th International Symposium on Malay / Indonesian Linguistics*, Jakarta, Indonesia, 26 July 2000.

Tadmor, Uri. "The Acquisition of Pronouns in Jakarta Indonesian", in U. Tadmor (ed.), *Studies in the Acquisition of Malay / Indonesian*, *NUSA Linguistic Studies in Indonesian and Languages in Indonesia*, Jakarta: Universitas Atma Jaya, (to appear).

INSA GÜLZOW AND NATALIA GAGARINA

ANALYTICAL AND SYNTHETIC VERB CONSTRUCTIONS IN RUSSIAN AND ENGLISH CHILD LANGUAGE*

Abstract. This paper deals with the differences in the early acquisition of verb constructions by two L1 Russian-speaking and two L1 English-speaking children and compares the emergence of finite and non-finite verb forms in the first six months of early verb production. As an analysis of the early verb forms will show, large language-specific differences exist with regard to the amount of finite vs. non-finite verb forms in the children's early productions although children acquiring either language display a preference for infinitival and/or other non-finite verb forms for a short period of time. Given the notable differences between Russian and English with regard to the richness of the inflectional system and the predominantly synthetic versus the predominantly analytical nature of Russian versus English verb constructions, the paper aims to show to what extent language-specific properties including the frequency of linguistic elements may contribute to the acquisition of different verbal constructions.

1. INTRODUCTION

Children acquiring Russian, English and other languages such as French, German, Dutch or Danish behave similarly in that they show a preference for using infinitival or base forms[1] during early stages of language acquisition (e.g. Rizzi 1994, Wexler 1994, Hoekstra and Hyams 1998, Wijnen et al. 2001). With regard to the emergence of finite verb forms and paradigmatic contrasts, on the other hand, it has been argued that a rich inflectional system (with little syncretism) facilitates the acquisition of finite verb forms (cf. Bittner et al. 2003). For children acquiring languages with a rich inflectional verb morphology such as Russian, Finnish or Estonian, contrastive verb forms are documented soon after the onset of verb production (e.g. Kiebzak-Mandera et al. 1997, Gagarina 2003, Laalo 2000, Vihman and Vija, this volume) while paradigmatic contrasts in the verb productions of children acquiring languages with a poorer verb morphology like English or French are attested much later (e.g. Gülzow 2003, Kilani-Schoch et al. 1997). In this paper we discuss some of the issues that are raised by comparing the emergence of non-finite and finite verb forms in the language of children acquiring Russian or English. We address, particularly, the question of what the accuracy of the children's use of finite verb forms (versus the inaccuracy of the children when producing verbal utterances with non-finite verb forms) can tell us about the relationship between the children's input and language development.

N. Gagarina and I. Gülzow (eds.), The Acquisition of Verbs and their Grammar: The Effect of Particular Languages, 229–259.

Discussions on how non-finite and finite verb forms are used at early stages of language development have called attention to the impact of a given target language's structure and patterns in the input. The production of so-called 'root infinitives', for instance, has been related to a variety of structural properties: Wexler (1994) argues for a connection between 'root infinitives' and the pro-drop factor; Hoekstra and Hyams (1998) relate it to the morphological richness of languages with verbal marking of number; Czinglar et al. (this volume), comparing English, German and Croatian acquisition data, argue for a connection between 'root infinitives' and copula omission. Contrary to approaches that relate the children's behavior to the characteristics of an internal grammar, a recent approach investigating Dutch children's verb development stresses the importance of salient elements in the input (Wijnen et al. 2001). Wijnen et al.'s approach focuses on the idea that children's immature language is a function of what is salient and that, due to its salience, an element can have a greater impact at a given stage of language development. Wijnen et al. argue that while this leads the children to produce non-target utterances with a non-finite form as the only verbal element, the children's behavior need not necessarily be interpreted as resulting from a deficiency of their early verb grammar.

It is known that at early stages of language development, children are sensitive to non-finite verb forms in analytical constructions and that they produce verbal utterances with only a single verbal element. An important aspect of such 'verb behavior' is the concept of salience. Salience is often defined in terms of the function of items in the input (e.g., content vs. functional categories; morphological, semantic, syntactic factors) and/or in terms of perceptual and acoustic properties. These factors facilitating the detection of a form in the input must be complemented by the uptake of a form by the language learner (cf. Gagarina 2005). Perceptual salience has been noted to play an important role both in the general cognitive development of children and, more specifically, in their linguistic development (cf. Golinkoff and Hirsh-Pasek 1996). While frequency effects are attested for a variety of linguistic forms (cf. Demuth 1989, Stromswold 1995), Naigles and Hoff-Ginsberg (1998) have noted for the acquisition of English verbs that the order of acquisition correlates with the frequency of verbs in the input, Vihman and Vija (this volume) show that the frequency of the Estonian form *on* 'is' in the early productions of one child mirrors the adult proportions. This effect is even greater if a given linguistic form occurs in sentence-final position (cf. also Slobin 1973, Shady and Gerken 1999), a position in which the linguistic element often receives prosodic stress. Gagarina (2002), who investigated the non-target use of infinitives in child Russian and the use of infinitives in the input, considered a conglomerate of features contributing to the degree of salience: acoustic features (intensity

and duration of the final palatalised [*t'*] facilitate perception); syntacitic features (stable final position in analytical constructions favors comprehension and production), and morphological properties (the lack of the inflectional suffixes favors the possibility of overgeneralized production).

While the majority of children are found to use mostly target-like finite verb forms in early synthetic verb constructions, the use of an infinitival or base form without a further verbal element is in most instances non-target-like.[2] For Russian, the non-target use of an infinitive with no further verbal element has been noted by various researchers since Gvozdev's (1949/1961) study (e.g. Snyder and Bar-Shalom 1998, Avrutin 1999, Brun et al. 1999). While English is also generally considered to be a language exemplifying the root infinitive phenomenon (e.g. Wexler 1994, Rizzi 1994, Hoekstra and Hyams 1998), opinions vary with regard to what kind of utterances are interpreted as instantiating the phenomenon. Wexler (1994) cites only those cases from English child language in which the verb is missing the 3S *–s*. Hoekstra and Hyams (1998) arrive at a much higher frequency estimate, ranging from 75% to 81% of all verb productions in early child language, including in their analysis all instances of a 'bare V construction' which correspond to verbs in the base form in our analysis.

We will argue here that the children's use of both infinitival or base forms without further verbal element and synthetic verb forms may be regarded as primarily reflecting the way in which the children interact with the structure of their target language as presented to them by the input. The absence of inflectional *–s* in examples like (1) and (2), for instance, need not necessarily be regarded as an instantiation of the child's failure to produce an appropriate inflection but may instead be interpreted as a reflection of the fact that in the target language 3S subjects may also occur in periphrastic verb constructions involving the base forms of the verb (it is possible to interpret these examples in terms of the child's awareness of a target verbal structure, see (3) and (4)).[3]

(1) Eve sit floor.[4]
(2) Cowboy Jesus wear boots.
(3) Can Eve sit on the floor?
(4) Cowboy Jesus should wear boots.

This is not to say that children producing utterances like (1) and (2) are in all cases aiming at periphrastic structures like (3) or (4). In this respect, our approach differs from positions that argue for the omission of (modal) auxiliaries in such cases (e.g. Klein 1974, Ingram and Thompson 1996). Rather, the children's use of a non-target base form in a phase of language development when their utterances mainly contain only one verbal element seems to reflect some basic knowledge about which verb forms are often

used in the language addressed to them. Our position is in accordance with accounts that argue for a strong influence of patterns in the target language. Contrary to those accounts that propose an immaturity of the children's grammatical competence when producing utterances with 'root infinitives', we will argue that children mostly use those main verb forms that they hear in the input (and that have high salience) and that their choice of a certain verb form is largely unaffected by the form's underlying grammar.

In the present paper we aim to show how the (main) verbs of a particular language's analytical and synthetic verb construction are mirrored in children's early verb productions. Our analyses are guided by the two basic assumptions that first, early verb production interacts with more general factors influencing children's language development such as the communicative interests which are characteristic of early child discourse and the complexity of verbal structures (cf. Tomasello 2003). Second, we believe that the salience of a verb form (including the frequencies of a certain finite or non-finite verb form in the input) will have an impact on that form's acquisition. We will show that in those cases in which Russian and English verb forms are similar with regard to communicative relevance, complexity, and salience, the interaction with the structure of the target language input results in similar occurrences of verb forms in the children's data.

The paper is organized as follows: section 2 and section 3 are devoted to a description of the two target languages' verb morphology and a brief description of the data used in the analysis. An analysis of the children's use of verb forms is carried out in section 4 to single out similar tendencies and divergences in the children's use of verb forms. The similarities and differences will be discussed in section 5 and section 6 as an effect of the interaction of the target language's structure and more general influences on language development. The conclusion in section 6 also refers briefly to children acquiring other languages such as Dutch, French or Spanish.

2. VERB MORPHOLOGY IN THE TARGET LANGUAGES

2.1. Verb Morphology in Russian

Russian is known to have rich verb morphology. The three tenses, past, present and future make a distinction between perfective (PFV) and imperfective (IPFV) aspect with PFV present being absent from the system although its formal representation realises the PFV future (Table 1). The majority of verbs belong to one of the two aspects and possess a counterpart. Forms of IPFV in the present and of PFV in the future (and the auxiliary *byt'* 'to be' in the analytical future with IPFV) are marked for person and

number while there is no person distinction in the past. Only singular past forms of both aspects are marked for gender. Table 1 presents the verb forms of the verb *menja-t'*–IPFV/*pomenja-t'*–PFV 'to change' in active verb constructions:[5]

Table 1. Menja-t'–IPF/pomenja-t'–PF 'to change'

Tense	*Number*	*Person (gender for past)*	*IPFV/PFV lexeme*
PAST	SG	MASC	(po)menja-l
		FEM	(po)menja-l-a
		NEUT	(po)menja-l-o
	PL		(po)menja-l-i
PRES	SG	1	menjaj-u
		2	menjaj-eš'
		3	menjaj-et
	PL	1	menjaj-em
		2	menjaj-ete
		3	menjaj-ut
FUT SYNTH. (PFV)	SG	1	pomenjaj-u
		2	pomenjaj-eš'
		3	pomenjaj-et
	PL	1	pomenjaj-em
		2	pomenjaj-ete
		3	pomenjaj-ut
FUT ANAL. (IPFV)			*byt'* 'to be' (1.2.3. SG or PL) + infinitive *menja-t'*

As Table 1 illustrates, the majority of forms in the verbal paradigm are synthetic. There are two types of analytical constructions (ACs) in Russian: with and without a finite verb form. ACs with a finite verb form include the auxiliary *byt'* 'to be', or phasal verbs denoting beginning, continuation or termination of an action, or the modal verb *hotet'* 'want' or any other finite verb and PFV and/or IPFV infinitive. An AC may also contain a non-finite modal element, i.e. temporal or modal predicatives, (mis)advising or (dis)allowing the performance of an action, e.g. *pora est'* 'it is time to eat', *mozhno chitat'* 'it is allowed/possible to read' or modal adjectives such as *nuzhno spat'* 'it is necessary to sleep'. Other types of sentences with infinitives are negations and impersonal constructions like *nechego myt'* 'there is nothing to wash', *emu est' chto nadet'* 'he has something to put on'.

One important characteristic of the inflectional system in Russian is the presence of two bases for each verb. The open base is usually related to the

stem and ends in a vowel, e.g., *smotr'e-* 'look', *menja-* 'change' (see Table 1); past tense forms are derived from the open base. The closed base is related to the root and ends in a consonant: *smotr'-*, *menjaj-*. Imperative singular forms are identical with the closed base and the present/future forms in the indicative and plural imperative are constructed from this base by attachment of inflectional endings.

2.2. Verb Morphology in English

In comparison to Russian, English is a predominantly analytical language which has lost most elements of its once rich inflectional system. Verb morphology is reduced to a few suffixes marking the categories person, number, tense and aspect and all inflections display a large amount of syncretism, including formal overlap with the marking of nominal categories such as the genitive case and the plural in the case of inflectional *-s*. In the present paper, the synthetic versus the analytical nature of the Russian versus the English verbal system will be the focus of interest. Children learning Russian are confronted with a majority of finite lexical verb forms while children learning English receive input with a majority of non-finite lexical verb forms that are part of an AC or are identical with the verb forms used in synthetic constructions. Judged on the basis of their formal properties alone, most English verb forms cannot be associated unambiguously with only one grammatical category.

In the absence of inflectional suffixes in English the categories person and number are mainly realized by auxiliaries and nominal elements such as personal pronouns. In analytical verb constructions English auxiliaries realize person and number distinctions along with tense, mood and voice. The morphology of English auxiliaries is opaque and here too a number of syncretic forms exist. Table 2 and Table 3 compare English PFV and IPFV verbal constructions in active voice to the respective verbal constructions in Russian.

Table 2. Analytical verb constructions in English in comparison to Russian

English		Russian	
PAST PERF/	had changed	PAST PFV	pomenja-l/-la/-lo/-li
- PROG	had been changing	PAST IPFV	menja-l/-la/-lo/-li
PRES PERF/	have/has …ed	PAST PFV	pomenja-l/-la/-lo/-li
- PROG	have/has been …ing	PAST IPFV	menja-l/-la/-lo/-li
PAST PROG	was/were …ing	PAST IPFV	menja-l/-la/-lo/-li
PRES PROG	am/are/is …ing	PRES IPFV	menja-u/-...
FUT (COND) I/-	will(would) …/	PRES PFV	pomenjaj-u/-...

PROG	will(would) be …ing	FUT IPFV	bud-u/- … + INF
FUT (COND)II/ PROG	-will (would) have …ed will (would) have been …ing		

Table 3. Synthetic verb constructions in English in comparison to Russian

English		*Russian*	
PAST	changed	PAST PFV (possible PAST IPFV)	pomenja-l/-…
PRES	change/changes	PRES IPFV	menjaj-u/-…

As Table 2 and Table 3 illustrate, except for the simple present and the simple past, all verbal constructions in English are analytical in nature.[6] An important point to mention is that besides the analytical constructions as listed in Table 2 and Table 3, English makes use of another kind of periphrastic verb phrase in negations and questions such as *She doesn't like grapes* or *Does she want some more?* In negations and questions, the otherwise non-analytical verb constructions simple present and simple past are also constructed periphrastically. With regard to the claim that the use of infinitives or base forms may be the result of omitting the auxiliary of a modal construction (e.g. Ingram and Thompson 1996), it is important to note that in English a common analytical construction involving the base form of the verb is not modal.

3. DATA DESCRIPTION

3.1. Russian

For this study we analyzed the longitudinal data of two L1-speaking Russian children within six months of the onset of verb production.[7] Both children grew up in monolingual middle-class families in St. Petersburg.[8] The children were systematically recorded and/or video-taped between one and six hours a month from the onset of speech until the age of three. The mean length of the recordings per month is about 150 minutes (2.5 hours). Table 4 gives the age of the children at the time of recordings, the relation of utterances with verbs (henceforth, verb utterances – VU)[9] to the number of all analyzed utterances (in percentages), and the mean length of utterance in words (MLU).

Table 4. VU and MLU (in words) by Vanja and Liza

Vanja				*Liza*			
AGE	VU	N	MLU	AGE	VU	N	MLU
2;1	5%	49	1.364	1;9	18%	147	1.055
2;2	14%	177	1.892	1;10	23%	86	1.191
2;3	33%	454	2.087	1;11	23%	69	1.199
2;4	35%	416	2.077	2;0	23%	64	1.688
2;5	39%	484	1.922	2;1	33%	92	2.249
2;6	44%	415	2.107	2;2	22%	65	2.636

3.2. English

For English we draw on the longitudinal data of two children included in the Brown corpus, Eve and Sarah (Brown 1973), as documented in the CHILDES database (cf. MacWhinney and Snow 1990, MacWhinney 1995). In comparison to the Russian children, the number of recorded utterances is largest in Eve's data and lowest in Sarah's data. To make the databases more comparable, only one recording per month for Eve and all available recordings per month for Sarah were used in the analysis.

Table 5. VU and MLU (in words) by Eve and Sarah

Eve				*Sarah*			
AGE	VU	N	MLU	AGE	VU	N	mean MLU
1;6	23%	162	1.537	2;3	18%	76	1.534
1;7	34%	82	1.946	2;4	18%	91	1.652
1;8	34%	233	1.988	2;5	20%	83	1.752
1;9	34%	177	2.310	2;6	22%	97	1.565
1;10	51%	255	2.938	2;7	22%	121	1.857
1;11	47%	206	2.830	2;8	32%	166	1.946

4. FINDINGS

Given the differences in the verb structures that the children learning the two languages hear with regard to the relative amount of synthetic versus analytic verbal structures, slightly different categories are used in the analysis. The verb forms that appear in the Russian data are differentiated with regard to five categories: infinitives (non-target use) such as *myt'* 'wash' and infinitives in target-like ACs such as *davaj chinit'* 'let us repair', synthetic verb forms such

as *slomal* 'break-PFV.PAST.SG.MASC' or *stoit* 'stand-IPFV.PRES.3S' and imperatives *risuj* 'paint', *daj* 'give'. Ambiguous and unanalyzable finite forms as well as a few agreement errors appear in the 'other' category. For the English children, verb forms represent one of the following five categories: past participles such as *gone*, past forms like *went* or *pulled*, base forms like *write*, 3S verbs like *wants* and verb forms ending in *–ing* such as *coming*. Verb forms that could not be categorized appear in the 'other' category.

4.1. Russian

The proportionate occurrence of verb forms in two children and in the language of their caretakers is exhibited in Figures 1 and 2. Among the five categories represented in the distribution of all verb forms, the second category basically stands out against the remaining four, since it shows the number of ACs within all VU (the four others represent the different types of (non-)finite synthetic forms). Since the emergence and stable production of AC is in inverse proportion with the non-target use of infinitives we describe all of these categories in one section.

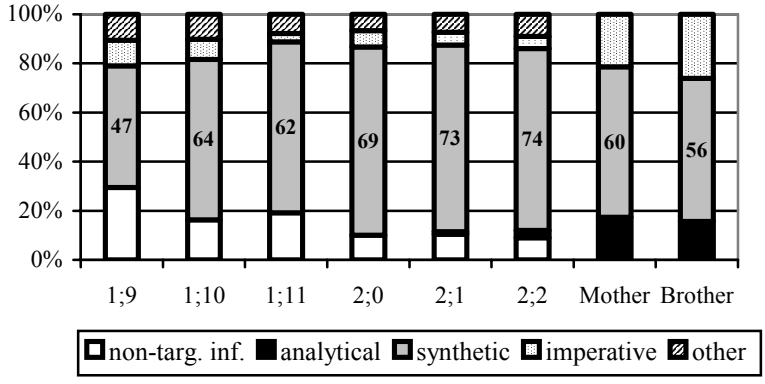


Figure 1. Liza's use of verb forms

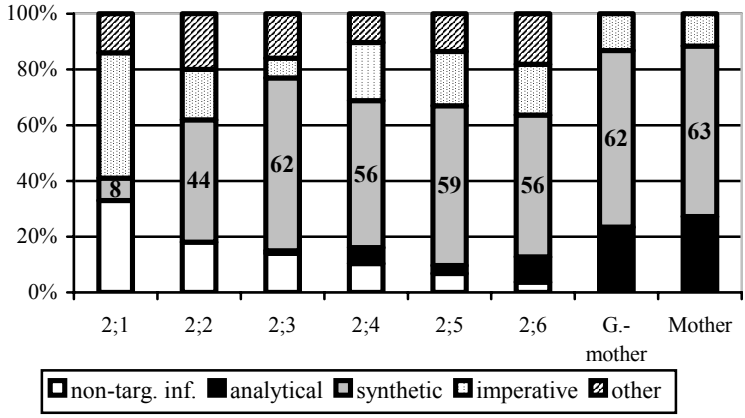


Figure 2. Vanja's use of verb forms

4.1.1. Infinitives and analytical constructions

The non-target use of infinitives includes the children's use of verb forms that constitute the lexical member of ACs in adult Russian, see (5) and (7).

(5) *myt'* (Liza 1;8 wants to wash teddy bear)
wash-INF
'(I) want to wash'

(6) *spat'* *bruu* (for *cars* Vanja 2;2a uses onomatopoeia *bruu*)
sleep-INF cars
'the cars are sleeping'

(7) *lomat'* *koleso* (Vanja 2;2b)
break-INF wheel-ACC
'I'm breaking the wheel'

Compared to the adult data, both children start with a high number of non-target infinitives: about one third of all VU. However, the number of these infinitives decreases steeply one month after their first occurrence: it drops by more than 10% percent and for the next two months remains between 14% -17%. The next decrease of another 10% is observed in L. within one month and in V. within two months. This second steep decrease leads to the full disappearance of infinitives within the following 4-8 months (this period is not analyzed here). Two-four months before full disappearance, infinitives in all three children occur only very sporadically (two – four

instances) (for details see Gagarina 2002). When the proportionate representation of infinitives in children's speech drops to 5%-6%, ACs begin to develop. This change is especially evident in Figure 2, which shows that at age 2;6 Vanja abandons the use of infinitives after ACs have been established in his system. In the non-target use of infinitives we find the largest discrepancy between the input and the children's speech only for a short period of time.

Both types of ACs, with and without a finite element, occur two-four months[10] after the onset of verb production: the interval between the onset of verb production and the emergence of ACs is bigger in Liza's data. This difference is connected with the slow development of the production of multi-component sentences in Liza (cf. MLU in Table 4, see also Gagarina 2003). Early ACs in both children do not differ with respect to the finiteness of the verbal elements they exhibit. The first and the only two instances of AC in Liza's data contain an infinitive with the finite modal verb (8) or the verbal particle (9) (the regular use of AC starts at 2;5).

(8)	*budu*	*slushat'*	(Liza 2;1)
	be-1S	hear-INF	
	'will hear'		
(9)	*davaj*	*poigraem*	(Liza 2;2)
	let us-IMP	play-PFV:TRANS-FUT:1P	

Vanja's early ACs contain predominantly utterances like (10) and (11), with the imperatives *davaj* 'let us!', *pojdjom* 'let us go' or non-finite predicates such as *pora* 'it is time', *nado* 'it is necessary' and an infinitive:

(10)	*chinit'*	*koleso*	*davaj*	(Vanja 2;3)
	repair-INF	wheel-ACC	let us-IMP	
	'let's repair the wheel'			
(11)	*pojdjom*		*igrat'*	(Vanja 2;4)
	go-PFV:TRANS-FUT:2P		play-INF	
	'let's go playing'			

A few instances of ACs with a finite element (auxiliary be-1S or the modal verb want-1S) are documented in Vanja already at 2;3, example (12), but the regular use of sentences as in (13) develops only towards the end of the analyzed period:

(12)	*budu*	*pryg*	(Vanja 2;3)
	be-1S	jump-ONOM	
	'will pryg (jump)'		

(13) *sobirat'* *budu* (Vanja 2;4)
collect-IPFV:TRANS-INF be-1S
'will collect'

4.1.2. Synthetic finite forms

VUs containing one finite verb (synthetic verb constructions) are mainly represented by PFV past tense forms and IPFV (3S and 3P) present forms and constitute the largest category in the data of both children from the onset of verb production. The only exception is documented in Vanja's data at age 2;1, when utterances with this type of verb forms constitute 8% of the total number of VU (cf. the number of correctly-used finite forms in two other Russian-speaking children, Kiebzak-Mandera, this volume). In Liza's data the number of finite forms is high from the onset of verb production and increases from almost 50% to over 70% in the following months. The proportional representation of synthetic verb forms in Liza's data is somewhat above the adult's from age 2;0 onwards. In Vanja's case we observe a slightly different trend. At the beginning, finite verb forms are very rare. At the age of 2;1 only four instances occur: *spit* 'sleep-PRV:3S:PRES' two times and the two past forms of *upast'* 'fell down' 'MASC:SG' and 'FEM:SG'. Towards the end of the period analyzed, synthetic verbs forms are represented by 50% of all instances. Contrary to Liza, their number stays a little below the proportionate representation of synthetic verb forms in the adult data.

4.1.3. Imperatives

In the first months of the analyzed period, both Liza and Vanja produce a larger amount of imperatives, 10% and 45% respectively, than they do in later sessions. However, the number of imperatives in Liza's data generally remains lower than in Vanja's data: a mean of 6% vs. 20%. The trend in Vanja's data is somewhat different in that although proportionate representation of imperatives decreases, in most later sessions it stays above the rate of imperatives in both Vanja's grandmother's (13%) and his mother's (12%) data. As the use of imperatives has a strong social and discourse component, it is not surprising that the two children differ in the use of this verb form. Liza expresses fewer demands and wishes, since she is generally permitted to perform her actions and there is not much need for her to produce imperatives. Contrary to Liza who tends to describe events, Vanja expresses his wishes more actively from the beginning of the recordings. He is cared for at his grandmother's home and frequently uses demands to obtain

desired objects. Besides, he wants adults to perform certain actions for him, like drawing a picture, building a house, etc., more often than Liza does.

4.1.4. Summary

The general trend to be observed in both Russian children is that soon after the onset of verb production the relative amount of different lexical verb forms basically mirrors that of the children's input. Just like adult Russian speakers, the children mainly produce synthetic verb forms and make few agreement mistakes. Target-like analytical constructions appear only sporadically towards the end of the analyzed period in Lisa but are more established in Vanja. In the earliest months analyzed, non-target infinitives and imperatives are over-represented in the children's data in comparison to the later recordings.

4.2. Verb forms: the English children

Figure 3 shows that analyzed on a purely formal basis, the distribution of the verb forms as documented in Eve's utterances basically mirrors the mother's use of verb forms. Only in the first two months of the study is the ratio of Eve's base forms higher in comparison with both her mother's and her own usage in later recordings. Although the number of instances in which Eve uses the base form is generally higher than the number of instances in which the Russian children use the bare infinitive, it is interesting to note that compared to the later recordings and similar to the Russian children, Eve produces a surplus of about 10-15% base forms in the first two months of the analysis. Regular uses of the present participle are documented in Eve's data from 1;8 onwards, with a mean representation of 16% in the last four months of the analysis. As in her mother's data, past participles are produced in very low numbers, with a maximum of 7% in the second month of the analysis. Inflected synthetic forms, 3S present forms and past forms, are produced only in a few instances by either Eve or her mother. While past tense froms are present in Eve's data from the beginning, only in the last month of the analysis do the first 3S present tense forms appear.

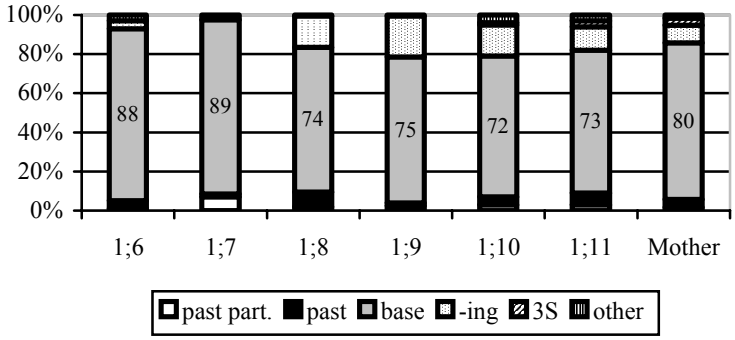


Figure 3. Eve's use of verb forms

In Figure 4, Sarah's and her mother's use of verb forms is illustrated. As with Eve, the distribution of verb forms in Sarah's utterances basically mirrors the distribution of verb forms in her mother's speech. At first sight the major difference to Figure 3 is that Sarah and her mother use the past participle in a larger amount of instances. As most of the instances involve rather formulaic constructions such as *(it's) all gone* or *(you've) got*, it is possible that the proportion for the past participle as documented in Figure 4 is the result of a preference for a certain verb phrase rather than the more frequent use of a productive verbal construction. Despite this, it seems worth noting that in the first month of the analysis, the ratio of the past participle is highest in Sarah's data, with a representation of 32%. The ratio decreases to 13% at age 2;4 and to 14% at age 2;5 after which this verb form almost vanishes from the data. With regard to Sarah's use of the base form, Figure 4 is similar to Figure 3 in that the largest amount of both Sarah's and her mother's utterances involves this form of the verb. The present participle is present from the beginning, but is represented both in Sarah's and in her mother's data by low numbers, ranging between 4% and 8%. As in Eve's data, inflected synthetic verb forms are rare, past forms are documented from the beginning with a maximum representation of 14% at age 2;6, and 3S present verb forms are hardly documented at all, the only instances being recorded at 2;6 and 2;7.

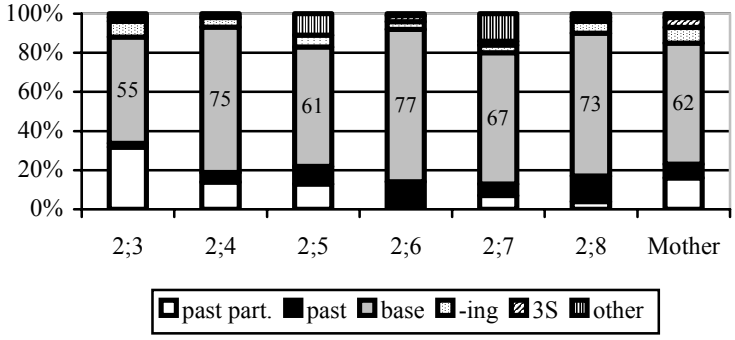


Figure 4. Sarah's use of verb forms

4.2.1. The base form

The most frequently used verb form in the English children's data is the base form. As already mentioned, the fact that Sarah's use of the base form is generally at a lower level than Eve's seems to reflect the distribution of verb forms in her mother's speech. After the first two months of analysis, in which about 90% of Eve's utterances contain the base form of a verb, the percentage decreases to about 75% in the following months. At 1;8 Eve's use of verbs ending in *–ing* is firmly established in her speech, which also seems to be the main reason for the decrease in her use of the base form. In comparison to the Russian children this is an interesting finding as it was shown that in Liza's and Vanja's data the decrease of infinitival and (base) imperative verb forms correlates with an increased representation of synthetic verb forms such as the IPFV present. The percentages of Sarah's use of the base form are somewhat different to those of Eve. After the first month of the analysis in which the percentage is low (55%), it increases to a mean of 70% in the following months. The most influential factor in Sarah's case seems to be the decrease of her use of the past participle from a level of 32% in the first month to 14% at age 2;4 and 13% at age 2;5. The main difference to Eve's data is that, rather than introducing a new form to her system, Sarah uses the past participle along with the base form from the beginning of her speech. The presence of a number of different verb forms in Sarah's data from the onset of the analysis may well be an effect of the later onset of verb production in Sarah who, at 2;3, has been exposed to her native language for a longer period than Eve whose onset of verb production is at 1;6. There is no straightforward parallel to Sarah's use of the past participle in Eve's or the Russian children's data, but it seems important to mention that like the base

form and the present participle in Eve's data, the past participle is the non-finite member of an AC and is over-represented for a brief period both with regard to later recordings and with regard to Sarah's mother's use of this verb form.

The children's use of verb forms as illustrated in Figures 3 and 4 disregards whether or not the form is used as part of a target-like structure. As illustrated in the previous section, the base form can serve a variety of functions in adult English. Apart from occurring as the non-finite member of ACs, epistemic and other state verbs are used in simple present constructions and the base is also used as the singular and plural imperative in English.[11] Figures 5 and 6 illustrate the children's and their mother's use of the base form in target-like simple present and ACs. For the ACs, modal structures and other periphrastic structures with the auxiliary *do* are distinguished. In the 'other' category all other cases appear. For the children the 'other' category also includes all instances of a possible use of the base form as an imperative as it is in most cases impossible to decide if a base verb form is used in this or another function. Although there are always a number of unambiguous instances in the data of the children, in most cases base forms that are potentially used as an imperative can also be interpreted as comments. In Eve mother's data imperatives constitute 22% of all verb productions/uses. Sarah's mother uses the imperative in 24% of all her utterances with verbs. The mothers' use of imperatives is not included in Figure 5 and Figure 6.

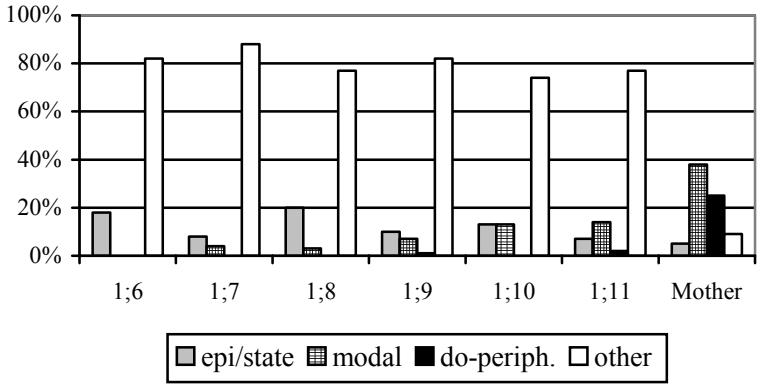


Figure 5. Eve's use of the base form

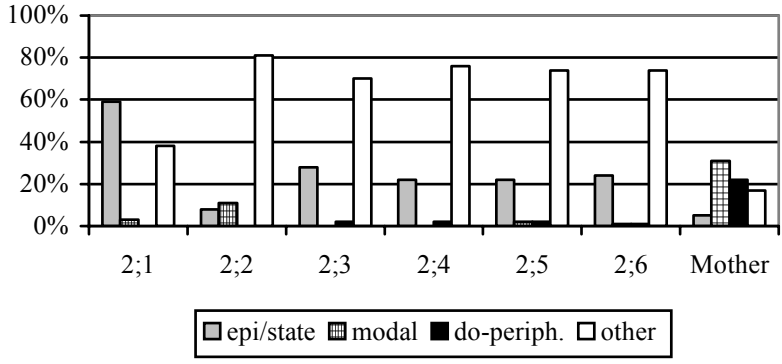


Figure 6. Sarah's use of the base form

Figures 5 and 6 show that Eve's mother and Sarah's mother are very similar in their distribution of the base form in different verbal structures. In the majority of instances the base form appears as part of an AC. These are mostly modal constructions (38% in Eve's mother's data, 31% in Sarah's mother's data), but a large number of base forms also occur as part of ACs with the auxiliary *do* (Eve's mother: 25%, Sarah's mother: 22%). In comparison to the Russian data as represented in Figures 1 and 2, these numbers translate into the following: 50% of all verbal utterances of Eve's mother are ACs with the base form. If differentiating between modal ACs and ACs with *do*, 30% of all verbal utterances consist of a modal plus a base form and another 20% of all verbal utterances consist of the auxiliary *do* and the base form. In the case of Sarah's mother, 18% of all verbal utterances consist of a modal and the base form and 13% consist of a form of the verb *do* and the base form. Taken together, 31% of all utterances of Sarah's mother are ACs with the base form. Epistemic verbs and other verbs denoting states are used only occasionally by either mother.

Contrary to the general distribution of lexical verb forms as illustrated in Figures 3 and 4, the children's use of the base form does not directly mirror the verbal structures that can be found in their mother's speech. For instance, the use of epistemic and other state verbs is represented in both children by a larger number of instances (Eve, mean of 13%, Sarah, 60% in the first month, in the following five months a mean of 20%) than in the speech of their mothers (Eve's mother: 5%, Sarah's mother: 5%). Target-like modal constructions appear more regularly in Eve's data only from 1;9 onwards; the first periphrastic constructions with the auxiliary *do* are not documented before the last month of the analysis. In Sarah's productions, target-like

analytical structures appear only in isolated cases. The largest number of instances in both children appears in the 'other' category. In Eve's and in Sarah's data this category is represented by about 80% of all instances throughout the period of analysis. Compared to the Russian children, the number of non-target infinitival/base forms seems to be greater and there is also no indication in Figures 5 and 6 that their use is about to decrease.

4.2.2. The present participle

Verbal constructions with the present participle are represented by 9% in Eve's mother's data and by 8% in Sarah's mother's data. While Eve begins to produce the present participle more regularly, from the third month of the analysis onwards (from 1;8), Sarah uses verb forms ending in *–ing* throughout the period studied at a mean of 4%. In Eve's data, both with regard to later recordings and with regard to her mother's speech, the use of verb forms ending in *–ing* is somewhat over-represented in the first months of its occurrence (1;8: 16%, 1;9: 21%, 1;10: 16%, 1;11: 12%). This is an interesting parallel to Eve's use of the base form, which shows a similar pattern. Although the base form is generally represented by a larger number of instances, the form is over-represented in the first months of the form's occurrence in comparison to the input and later recordings. Table 6 shows that the present participle also tends to appear as part of an AC without its auxiliary at early stages of verb production. In the majority of cases the present participle is produced in utterances without an overt subject so that correct agreement was difficult to assess. A couple of instances appeared in questions. The remainder were categorized with regard to the target-like realization of the verb phrase. In the majority of instances the present participle occurred as the only verbal constituent of the child's utterance (NP) throughout the period of analysis. Only in very isolated cases was agreement realized by the use of a cliticized auxiliary (NP'm/'re/'s) or a full auxiliary (NP am/are/is).

Table 6. Eve's use of the present participle

	1;6	*1;7*	*1;8*	*1;9*	*1;10*	*1;11*
no subject	1	2	12	14	15	4
questions	-	-	8	1	-	1
NP	2	-	6	20	19	16
NP'm/'re/'s	-	-	1	-	1	-
NP am/are/is	1	-	-	-	1	-
total	4	2	27	35	36	21

Although Sarah uses fewer verbal constructions with the present participle, the finding is similar in that during the six months of the analyzed period utterances with this verb form lack a subject, in most instances. In those utterances that do have a subject, verb forms ending with *–ing* are used with no other verbal constituent.

Table 7. Sarah's use of the present participle

	2;3	*2;4*	*2;5*	*2;6*	*2;7*	*2;8*
no subject	4	2	3	-	-	3
questions	-	-	-	1	-	1
NP	-	1	1	2	4	3
NP'm/'re/'s	-	-	-	-	-	-
NP am/are/is	-	-	-	-	-	-
total	4	3	4	3	4	7

4.2.3. The past participle

In the six months of the analysis Eve and Sarah produce a total of 24 and 47 past participles. There is only one instance in Eve's data in which the target utterance may be a present perfect construction. All other instances express states either aiming at copula constructions involving expressions such as *gone* (14) or *done* (15) and in later recordings expressions such as *broken* (16), or are examples of *have/has got* (17), which in adult language is also often elliptical leaving out the auxiliary (18). Only in two (Eve) or one (Sarah) example(s) do Eve and Sarah express a copula, see (19) and (20).

(14)	*my teddy bear gone*	(Sarah 2;1)
(15)	*all done nap*	(Eve 1;9)
(16)	*broken*	(Sarah 2;3)
(17)	*you got K(leen)ex?*	(Sarah 2;6)
(18)	*what you got?*	(Sarah mother)
(19)	*the dog is stuck*	(Eve 1;7)
(20)	*Bobo's gone*	(Sarah 2;1)

4.2.4. The past tense form

The overwhelming majority of past references in the two children's data consist of irregular past tense forms. In Eve's data regular past tense forms appear only towards the end of the period analyzed and in Sarah's data regular past reference is documented only in a few isolated cases. In line with studies that argue for a non-finite phase preceding a finite phase in English

verb development (Weist 1986), it seems that the regular morphological marking of verbs appears relatively late in English children's data, following a phase in which irregular past tense forms with a similar phonological structure have been documented for some months (cf. Gülzow 2000).

Table 8. Eve's and Sarah's use of past forms

Eve	*1;6*	*1;7*	*1;8*	*1;9*	*1;10*	*1;11*
irregular	4	1	11	2	5	8
regular	-	-	2	-	5	2
Sarah	*2;3*	*2;4*	*2;5*	*2;6*	*2;7*	*2;8*
irregular	1	3	4	10	6	16
regular	-	-	2	1	-	-

4.2.5. 3S forms

It is a well-documented fact that in English the target-like marking of 3S forms in the simple present emerges relatively late (e.g. Brown 1973). At first sight this finding looks like a contrast to the finding that in Russian present references marked for 3S are among the first verbal forms to appear. As already mentioned earlier, simple present 3S forms in English are considered habitual and generic while in Russian only IPFV present forms exist; the different function of the English form together with its rare occurrence in the adult data must be regarded as major influences on its role in the children's early verbal systems.

4.2.6. Summary

The similar general tendency for both Russian and English children is that the different verb form frequencies approach the adult representations in the course of the first six months of verb use. An exception to this pattern are non-finite forms like the infinitive and the imperative verb form in Russian and the base form and the present and the past participle in English. For these forms, brief periods of over-representation are documented here. Imperatives and infinitives are over-represented at the onset of verb production in the Russian children's data. The English children over-represent mainly the base form, but there is also some indication that other non-finite verb forms such as the present participle (Eve, cf. Figure 3) and the past participle (Sarah, cf. Figure 4) occur in a greater number of instances in some months. Even though English children produce more base forms than the Russian children produce infinitives in this period, the overproduction ofbase forms in comparison to later recordings (Eve) and in comparison to the input that they

receive (Eve and Sarah), does not exceed the level of surplus utterances with the infinitive in the Russian children's data.

With regard to the structure of target-like verb phrases, larger differences to the adult data could be observed. All children are late in their production of target-like ACs. This effect is especially strong for ACs with (non-modal) auxiliaries. Since for the Russian children this kind of construction is used for future reference, only isolated instances are documented in their data. For the English children, modal auxiliaries plus the base form of a verb are the first ACs to appear. Other ACs in which the auxiliary's meaning is less transparent are hardly documented in the period of the analysis. Present perfect constructions are absent from the English children's data, if the past participle is used, the target structure is in most cases a copula construction in which the past participle functions as an adjective. Like the present participle, it mainly occurs without auxiliary. Only in the latest analyzed months of Eve's data are isolated instances of the auxiliary *do* plus the base form documented.

5. DISCUSSION

As the analyses presented in the previous sections have shown, the Russian and the English children in the study behave similarly in that their choice of a main verb form basically mirrors the input they receive. With regard to their target language, Russian children are much more accurate in their productions, as soon after the onset of verb use their utterances contain a finite verb form with correct agreement features. On the contrary, a large number of English children's early verbs are base forms or other non-finite verb forms without additional verbal element; agreement features are not expressed. Despite these differences in the Russian and the English children's early verb productions, all children display a tendency to use infinitival and other non-finite verb forms in greater frequencies than the input would predict for short periods of time. In the following, we will firstly discuss the implications of the finding that the children produced main verb forms in proportions that are in accordance with their input. Secondly, we will discuss the Russian and the English children's use of non-finite forms.

5.1. Main verb forms used by adults and children

The strong correlation between the proportions of verb forms in the child and in the adult data seems in large parts to be motivated by the relevance of a given verb construction in child discourse. Children produce verb forms that are part of relevant, and therefore frequently attested, verbal constructions in child-directed speech at an early point in language development. For instance,

when acquiring German as their first language, children produce verb forms that are used for present reference such as *badet* 'is bathing' and *malt* 'is drawing' before they systematically use verb forms such as *badete* 'was bathing' and *malte* 'was drawing' since past reference is more commonly made with analytical *habe gebadet* 'have been bathing' and *habe gemalt* 'have been drawing' (cf. Behrens 1993). For the present paper it is important to note that this finding is relevant not only for verb forms in utterances with a (single) finite verb form but also for the main verb forms in analytical verb constructions used in child-directed speech. Given the fact that children mainly produce utterances with only one verbal element at the early stages of language development, the main difference in the Russian and the English children's use of finite versus infinitival or base forms may be regarded as a consequence of the two target language's distinct structures. For the same reason – children's early use of relevant (single) main verb form - agreement is expressed at an earlier stage by the Russian than by the English children. As already mentioned earlier, the late appearance of 3S *–s* in English is therefore not a contradiction in comparison to the early synthetic forms in Russian children but mirrors the input in that simple present verb forms are generally represented by low frequencies in the input, that is, the habitual or generic use of the simple present has a function that is not highly relevant in a young child's discourse. With regard to a more advanced stage of language development, the late emergence of agreement in the English children is not only connected to the predominance of the base form in verbal utterances, but is also documented by the finding that non-modal auxiliaries which agree with the subject noun phrase are generally used later than modal auxiliaries which do not express agreement.

As we consider the children's choice of a verb form to be mainly motivated by the forms they hear in the speech addressed to them, agreement must not be regarded as a systematic feature of an underlying grammar. Rather, it can be interpreted as an accompanying feature of the early verb forms used by the children. Such an interpretation is in line with the finding that although Russian children establish paradigmatic contrasts early, almost no overgeneralizations are documented at this point in their language development, (cf. Gagarina 2003). Although the Russian children use a variety of contrasting inflectional verb forms, the very low number of overgeneralizations supports the position that at early stages of verb use the morphological structure of the forms remains unanalyzed. The finding that Russian children in the first months of verb production establish inflectional affixes in an order in which 3rd and 1st person marking precedes 2nd person marking, and in which singular marking precedes plural marking, can be correlated with the order in which English children make subject noun phrase reference in their early VUs. That is, the appearance and order of (present)

inflectional affixes in Russian is connected to a more general order that can be observed for the expression of referents, but need not be ascribed to a systematic difference in the children's verb grammars. In contrast to the Russian children, overgeneralizations are noted at the onset of a more regular use of inflectional affixes by the English children. English children mostly use irregular past forms before regular *–ed* is established and begins to occur in overgeneralizations. At about the same time, 3S *–s* appears more regularly in the language of most English children (cf. Brown 1973). In a more radical interpretation of these findings it even seems possible to argue that the English children might be more aware of inflectional affixes at some point of development, since inflected forms are the exception rather than the rule in the system that they acquire. Positions (e.g., Weist 1986) in which it is argued that English children pass from a non-finite to a finite phase, in which the few inflectional affixes of English arise in concert, would support such a claim (cf. Gülzow, 2003).

The idea that it is not grammatical representations in the children's early language that lead to large differences between the number of finite versus non-finite verb forms in Russian and English children's early verb use is also supported by the finding that the functional array that a verb form can serve in a particular language seems to have an impact on the children's VUs. For instance, the observation that in languages like Dutch most uses of the infinitive form by the children are modal, while in English not all uses of the base form can have a modal interpretation need not necessarily be connected to the language's morphology (cf. Hoekstra and Hyams 1998), but seems to mirror the structure of the learner's target language. Dutch periphrastic verb constructions involving the infinitive are mostly modal while in English they can be but need not be. If the adult's use of a certain verb form is multi-functional, it is not surprising that the children do not use it in a single function either.

5.2. The children's use of non-finite verb forms

When producing utterances with verbs, both the Russian and the English children choose verb forms in frequencies that soon approach the proportions found in the input. An exception to this trend is those months in which an over-representation of non-finite verb forms is documented. As already mentioned in the introduction, children's preference for infinitival verb forms at early stages of verb production has been referred to as the 'root infinitive' phenomenon. In these discussions, the concept of salience has played a significant role. On the basic assumption that salient elements can have a greater impact on language development for a given period of time, we will compare our findings in the Russian and the English data and extend the

argument to the children's use of non-infinitival yet nonfinite verb forms such as the English participles.

Properties such as frequency, prosodic stress, sentential position and lexical/conceptual transparency have been recognized as contributing to a given verb form's salience. Wijnen et al. (2001) showed that the positional and perceptual salience of the infinitive in Dutch ACs leads to a strong preference for this verb form in the early productions of Dutch children while the appropriate auxiliary remains absent for a relatively long period. Although Dutch is a verb second language, the main verb of periphrastic verb constructions appears sentence-finally, see (21); example adapted from Wijnen et al. (2001):

(21) *Maurits wil Marilène zoenen.*
Maurice want-3S Marilyn kiss-INF
'Maurice wants to kiss Marilyn'

Various studies have demonstrated that elements in sentence-final position exhibit a developmental advantage (they are acquired more easily and at earlier stages of language development: Slobin 1973, Shady and Gerken 1999). The finding is sometimes attributed to the prosodic stress associated with this sentential position. As stated in section 2, the non-finite members of ACs in Russian share positional and perceptual salience with the infinitival form in Dutch. While word order in Russian is fairly flexible, sentences with ACs in which the main verb occurs as an infinitive exhibit a more stable word order with the infinitive in sentence-final position. With regard to word order, the position of the non-finite element in English is quite similar to Russian ACs, as the non-finite verb form of an AC occurs in sentence-final position if there are no further constituents, see (22) and (24). Since it directly follows the conjugated verb, there are some cases in which the non-finite verb form will not be in sentence-final position, see (23) and (25).

(22) *I will run /I like swimming.*
(23) *I will run for two hours each day/ I like swimming in the ocean.*
(24) *Ja budu begat'*
I be-1S run-INF (multidirectional)
'I will run'
(25) *Ja budu begat' dva časa každyj den'*
I be-1S run-INF two hours each day
'I will run for two hours each day'

In terms of the prosodic stress associated with sentence-final position, it can be argued that English non-finite verb forms are similar to Russian infinitival verb forms if they occur in sentence-final position, see (22) and (24). The positional and prosodic salience effect of the non-finite verb forms in (23) may be regarded as similar to that of the non-finite verb forms in (25); from a sentence-internal perspective the non-finite verb in both (23) and (25) is more salient than the auxiliary.

In light of the similar degree of salience of the English and Russian non-finite forms it may seem difficult to explain the finding that, despite a generally greater amount of instances with the base form in the English children's data, the Russian children over-represent the bare infinitive in the first months of the analysis to a much greater extent. In the first two months of the study, Liza produces 130% more VUs with an infinitival form in comparison to her mean number of utterances with an infinitive in later recordings. In the first two months of the study Liza also produces 88% more utterances with an infinitive than her mother and brother. For Vanja the surplus of infinitives in comparison to his later recordings amounts to 120% and in comparison to his grandmother and mother it amounts to 32%. Quite differently, the mean percentage of Eve's utterances with a base form at age 1;6 and 1;7 is only 11% above that of her mother; Sarah's mean of utterances with the base form between the ages of 2;4 and 2;8 is 15% above that of her mother. In comparison to later recordings, Eve produces only 20% more utterances with the base form in the first two months of the study. Although it is logically impossible that the English children produce a surplus of utterances with the base form at as high a level as the Russian children produce VUs with the infinitive, it seems worth noting that the impact of the non-finite form in Russian is strong, given a generally lower representation of that verb form in the target language. This finding supports the claim that a stable sentence-final position of an element has more influence than the element's frequency. The other non-finite forms in the English children's data are represented by a surplus of 77% (Eve's use of the present participle between the ages of 1;8 and 1;11) and 100% (Sarah's use of the past participle at age 2;3) in comparison to their mother's use of the form. As no within-child comparison is possible for the use of these forms, it is difficult to say whether the impact of the increased salience is comparable to that of the base form, but we interpret this finding as a tendency in that direction. On the basis that both the present and the past participle in periphrastic verb constructions serve as the main predicate and are thus semantically, positionally and prosodically similar to the other non-finite verb forms in English, it does not seem surprising that these verb forms are over-represented in Eve's and Sarah's data for brief periods of time.

It has been argued that within a language's verbal system, different degrees of lexical or conceptual transparency exist (cf. Schlichting 1996, Wijnen et al. 2001). Eventive verbs or, more specifically, (transitive) action verbs denote activities that are perceptually delimited and salient. Schlichting (1996) has claimed that in Dutch adult speech, finite verb forms can be correlated with state verbs while infinitives primarily denote (transitive) activities. While it remains an open question whether the same kind of relationship exists between the Russian infinitive and other finite main verbs or whether in English the fact that the base form can represent both a synthetic and part of an analytical construction plays any important role, the degree of abstractness versus 'concreteness' of a verbal element's reference seems also relevant for the acquisition of different auxiliaries. Modals like Russian *nado*:ADJ:PRED (advises the action), *mozhno*:ADV:PRED (allows the action) or *nezl'ja*:NEG:PRED (disallows the action) and English *should/shouldn't* and *could/couldn't* contribute to sentence meaning in similar ways in that they make a statement about abilities, probabilities, permission, etc. In contrast, the (non-modal) auxiliaries of Russian and English like Russian *byt'* or English *do* and *be*, do not contribute to the sentence meaning in a comparable manner. For instance, the auxiliary *do* in English is used in periphrastic constructions like questions and negations without otherwise changing the semantic content of the predicate.

(26) a. I play the piano.
b. Do you play the piano?
c. Do not play the piano!

In line with the early production of lexically or conceptually highly transparent verbs (as well as pragmatically salient ones), the children in the study opted for modals with a relative high degree of lexical or conceptual transparency when producing full analytical constructions. For the English children, this finding is especially relevant as it explains why the majority of their verbal structures remain non-target-like for a longer period than do the verbal structures of the Russian children. As Figure 5 and Figure 6 illustrated (cf. section 3), periphrastic verb constructions with *do* and the base form are represented by almost as many instances in the input as modal constructions but hardly ever occur with the auxiliary in the first six months of verb production – while both children produce a number of modal constructions. The same observation was made for the English children's use of verbal constructions with the present and the past participle. Throughout the six months analyzed the participles were used without auxiliary in the overwhelming majority of instances.

6. CONCLUSION

To summarize, we hope to have shown that the contrasting findings of early production of target-like synthetic verb forms by Russian children versus the long persistence of a large number of non-target base form utterances in children acquiring English are not necessarily rooted in the grammatical representations underlying children's early verb systems. Rather, the way in which the structure of the target language interacts with more general developmental processes seems to determine the children's early choices of verb forms. For instance, when making present reference, the Russian children will in many cases choose an appropriate inflected form. The English children, on the other hand, mainly produce the base form of the verb, which can be used in simple present constructions but which mainly occurs in the input as part of an AC. We could observe that when verbs ending with the suffix *–ing* occur in the children's data they do so with no auxiliary, but they soon approach the adult standard with regard to frequency. It is therefore no contradiction that Russian children readily produce the 3S present inflection while English children only do so in isolated instances (see above).

The occurrence of the early verb forms in both Russian and English child language is determined by a set of factors including a given form's frequency in the input, its (prosodic and positional) salience and communicative relevance. Of these, the concept of salience seems to be of most importance as a way of accounting for the similarities noted in the children's use of non-finite verb forms. While Sarah's early verbal utterances seemed to be strongly influenced by a late onset of language production and generally slow progress in terms of the overall syntactic complexity of her utterances, the comparison of Eve and the two Russian children Liza and Vanja revealed that they behave similarly with respect to their use of the non-finite members of ACs.

It is clear that many of our findings confirm already established facts. What we hope to have shown is that the choice of a verb form by children acquiring languages with a primarily synthetic versus a primarily analytical verbal system may be more input-based than is often supposed. At early stages of language production children seem largely indifferent as to whether they produce a form that is finite or not. Irrespective of the fact that children are clearly sensitive to agreement features as documented in their target-like use of these forms, we could find no indication in the data that the choice of a finite versus a non-finite verb form is determined by anything but the way in which main verb forms relate to the verbal system.

A next step in the direction of isolating the language-specific versus the more general influences on early verb use would naturally involve the analysis of other languages. It has, for instance, been noted that French

exemplifies the root infinitive phenomenon while Spanish does not (Torrens 1995, Kilani-Schoch 2000). This finding has been connected to the absence versus the presence of the pro-drop factor in the two languages. Although the verbal systems of French and Spanish are in many ways very similar, the languages also differ with respect to their realization of non-finite verb forms in ACs. From our perspective, it would be an interesting question whether the fact that, unlike in French, it is possible in Spanish to attach a given pronoun to the non-finite member of an AC in any way contributes to the classification of Spanish as a non-'root infinitive' language. Another open question concerns the finding that, despite the fact that child-directed speech in Russian and in Dutch contains about 30% of ACs with an infinitival verb form, Russian and Dutch children behave differently in that Dutch children go through a phase in which they use the infinitival form almost exclusively (e.g. Wijnen and Verrips 1998) while Russian children do not. In sum, we hope to have shown differences and similarities in early children's verb production and the influence of the language specific structures and their frequencies on the early acquisition of verbs.

INSA GÜLZOW
Zentrum für Allgemeine Sprachwissenschaft (ZAS), Berlin

NATALIA GAGARINA
Zentrum für Allgemeine Sprachwissenschaft (ZAS), Berlin

7. NOTES

* We would like to express our gratitude to Marilyn Vihman who contributed to this paper and to the book with her suggestions, enthusiasm and encouragement. We also thank Philippa Cook for her help concerning formulation and stylistics. All mistakes are our own.

1 In the present paper the term 'infinitive' or 'infinitival (verb) form' refers to verb forms that carry an inflection; it is distinguished from the term 'base form' which is a verb form that can also be used as an infinitive, but is not inflected.

2 Lasser (2002) shows that adults also use infinitives without further verbal element in informal speech and discusses the children's use of this construction as a reflection of the adult's behavior.

3 Examples (1) and (2) are cited by Hoekstra and Hyams (1998) and were originally taken from Eve's data of the Brown corpus (CHILDES, Brown 1973).

4 We disagree with the claim that the target structure underlying the example in (1) is 'Eve sits on the floor'. Rather it seems that 'Eve is sitting on the floor' is more appropriate. As we do not argue that children's productions of the sort cited in (1) compete with identical utterances with 3rd person singular *-s* but reflect analytical verb constructions in which the base form of the verb is the target-like main verb form of the sentence we will not pursue this point further.

5 Passive verb constructions are analytical and consist of a reflexive conjugated verb with the postfix *–sja*, e.g. *dom stroitsja* 'house-nom build-IPFV.3.sg' *the house is built* and will not be discussed in detail here as they did not occur in the very early productions of either the Russian or the English children.

[6] This is also true for passive constructions which display the tense properties of the respective active construction in a finite realization of the verb *to be* combined with the main verb as a past participle, e.g. active: *He baked the cake* versus passive: *The cake was baked*.

[7] We exclude from the analysis the very early sets of data where utterances with verbs are below twenty instances. For Liza thus, the first two months of verb production with the number of verb tokens 1 and 13 were excluded.

[8] Liza is the second child in a family of linguists.

[9] 'N' refers to the actual number of utterances with verbs.

[10] To recollect, we did not calculate the first two months in Liza's data.

[11] The use of the base to denote habitual activities and events is included in the 'other' category. This function hardly occurred in the children's productions. In Sarah's mother's speech the use of the base form to describe routines was at one point in the recordings quite frequent and explains the high score of the 'other' category in her data.

8. REFERENCES

Avrutin, Sergey. *Development of the Syntax-Discourse Interface*. Dordrecht: Kluwer Academic Publishers, 1999.

Behrens, Heike. *Temporal reference in German child language.* PhD Dissertation, University of Amsterdam. 1993.

Bittner Dagmar, Wolfgang U. Dressler, and Marianne Kilani-Schoch (eds.), *Introduction: Cross-linguistic approaches to the acquisition of early verb inflection*. Berlin: Mouton de Cruyter, 2003. x-xxxvii.

Brun, Dina, Sergej Avrutin, and Marina Babyonyshev. Aspect and its Temporal Interpretation during the optional Infinitive Stage in Russian. Paper presented at *BUCLD 23: the 23rd annual Conference on Language Development*, Boston University 1999.

Brown, Roger. *A first language: the early stages*. Cambridge, MA: Harvard University Press, 1973.

Demuth, Katherine. "Maturation and the acquisition of the Sesotho passive." *Language* 65, 1989: 56–80.

Gagarina, Natalia. "Thoughts on Optional Infinitives (in Russian)." In Daniel Hole, Paul Law and Niina Zhang (eds.), *Linguistics by Heart. Webfest for Horst-Dieter Gasde*. Berlin: ZAS, <http://www.zas.gwz-berlin.de/mitarb/homepage/webfest/>, 2002: 1–23.

Gagarina, Natalia. "The early verb development and demarcation of stages in three Russian-speaking children." In Dagmar Bittner, Wolfgang U. Dressler, and Marianne Kilani-Schoch (eds.), *Cross-linguistic approaches to the acquisition of early verb inflection.* Berlin: Mouton de Cruyter, 2003: 131–170.

Gagarina, Natalia. "Salience as a factor in models of language acquisition". Paper presented at the Xth International Congress for the Study of Child Language, Berlin, 2005.

Gelderen, Van and Van der Meulen. *Root infinitives in Russian: Evidence from Acquisition*. Term paper, Leiden University, 1998.

Golinkoff M. Roberta and Kathy Hirsh-Pasek (eds.), *The origins of grammar: Evidence from early language comprehension*. Cambridge, MA: MIT Press, 1996.

Gülzow, Insa. "Early Development of Verb-constructions in an English-speaking Child." In Dagmar Bittner, Wolfgang U. Dressler, and Marianne Kilani-Schoch (eds.), First Verbs: On the way to Mini-paradigms. *ZAS Papers in Linguistics* 18, 2000: 39–52.

Gülzow, Insa. "Early development of verbal morphology in an English-speaking child." In Dagmar Bittner, Wolfgang U. Dressler, and Marianne Kilani-Schoch (eds.), *Cross-linguistic approaches to the acquisition of early verb inflection.* Berlin: Mouton de Cruyter, 2003: 205–238.

Gvozdev, Aleksandr N. *Formirovanie u rebenka grammaticheskogo stroja russkogo jazyka*. Moskva: APN RSFSR, 1949/1961.

Hoekstra, Teun and Nina Hyams. "Aspects of Root Infinitives." *Lingua* 106 (1998): 81–112.

Ingram, David and William Thompson. "Early syntactic acquisition in German: Evidence for the Modal Hypothesis." *Language* 72 (1996): 97–120.

Kiebzak-Mandera, Dorota, Magdalena Smoczyńska, and Ekaterina Protassova. "Acquisition of Russian verb morphology: the early stages." In Wolfgang U. Dressler (ed.) *Studies in Pre- and Protomorphology*. Wien: Verlag der Österreichischen Akademie der Wissenschaften, 1997: 101–114.

Kilani-Schoch, Marianne. "Early verb development of two French-speaking children". In Dagmar Bittner, Wolfgang U. Dressler, and Marianne Kilani-Schoch (eds.), First Verbs: On the way to Mini-paradigms. *ZAS Papers in Linguistics* 18, (2000): 79–98.

Kilani-Schoch, Marianne, Anna de Marco, Anastasia Christofidou, Maria Vassilakou, Ralf Vollmann, and Wolfgang U. Dressler. "On the demarcation of phases in early morphology acquisition in four languages." In Katarzyna Dziubalska-Kołaczyk (ed.), *Pre- and Protomorphology in Language Acquisition*. Papers and Studies in Contrastive Linguistics 33, Poznań: Adam Mickiewicz University, 1997: 15–32.

Klein, Raymond M. "Word order: Dutch children and their mothers." *Publications of the Institute of General Linguistics 9,* University of Amsterdam, 1974.

Laalo, Klaus. "Early verb development of two Finnish-speaking children: a preliminary approach to miniparadgms." In Dagmar Bittner, Wolfgang U. Dressler, and Marianne Kilani-Schoch (eds.), First Verbs: On the way to Mini-paradigms. *ZAS Papers in Linguistics* 18, (2000): 53–68.

MacWhinney, Brian and Catherine Snow. "The Child Language Data Exchange System: An update." *Journal of Child Language*, 17 (1990): 457–472.

MacWhiney, Brian. *The CHILDES Project: Tools for Analyzing Talk*. 2nd ed. Hillsdale, NJ: Erlbaum, 1995.

Naigles Letitia R. and Erika Hoff-Ginsberg. "Why are some verbs learned before other verbs?" *Journal of Child Language* 25 (1998): 95–120.

Rizzi, Luigi. "Some notes on linguistic theory and language development: the case of root infinitives." *Language Acquisition* 3 (1994): 371–393.

Schlichting, J.E.P.T. *Discovering syntax. An empirical study in Dutch language acquisition*. PhD Dissertation, University of Nijmegen, 1996.

Shady, Michele and Louann Gerken. "Grammatical and caregivers cues in early sentence comprehension." *Journal of Child Language* 26 (1999): 163–176.

Slobin, Dan I. "Cognitive prerequisites for the acquisition of grammar." In Charles A. Ferguson and Dan I. Slobin (eds.), *Studies of child language development*. New York: Holt, Rinehart and Winston, 1973: 175–208.

Snyder, Willyam and Eva Bar-Shalom. "Word Order, Finiteness, and Negation in Early Child Russian." Paper presented at *BUCLD 22: the 22nd annual Conference on Language Development*, Boston University 1998.

Stromswold, Karin. "The acquisition of subject and object Wh-questions." *Language Acquisition*

Tomasello, Michael. *Constructing a Language*. Harvard University Press. 2003.

Torrens, Vicenç. "The Acquisition of Inflection in Spanish and Catalan." In Carson T. Schuetze, Jennifer B. Ganger, and Kevin Broihier (eds.), *Papers on Language Processing and Acquisition*. Cambridge: MIT Working Papers in Linguistics, 1995: 451–472.

Weist, Richard M. "Tense and Aspect: Temporal systems in child language." In Paul Fletcher and Michael Garman (eds.), *Language Acquisition*. Cambridge: Cambridge University Press, 1986: 356–374.

Wexler, Kenneth. "Optional infinitives, verb movement and the economy of derivation in child grammar." In David Lightfoot and Norbert Hornstein (eds.), *Verb movement*. Cambridge: Cambridge University Press, 1994: 305–382.

Wijnen, Frank and Maaike Verrips. "The acquisition of Dutch syntax." In Steven Gillis and Annick De Houwer (eds.), *The acquisition of Dutch*. Amsterdam: Benjamins, 1998.

Wijnen, Frank, Masja Kempen, and Steven Gillis. "Root infinitives in Dutch early child language: an effect or an input?" *Journal of Child Language* 28 (2001): 629–660.

Part 5

Language-specific and learner-specific peculiarities in the development of verbs and their grammar

MARILYN MAY VIHMAN AND MAIGI VIJA

THE ACQUISITION OF VERBAL INFLECTION IN ESTONIAN

Two Case Studies

Abstract. The first six months of verb production is traced in two boys acquiring Estonian, based on both recordings and diary notes. Evidence of the emergence of contrast is identified and related to cumulative verb lexicon. Rate of verb learning is found to be variable whereas contrast emerges at a relatively constant rate for the two children. First contrastive use in context is illustrated for one child. The use of stems in lieu of obligatory inflections is described, as are phonological and morphological errors (errors of both omission and commission). The input frequency for all verb forms in the speech of one mother is analysed for tokens and types; only one exceptionally high frequency verb form appears to have had a direct effect on the child. The children are found to differ primarily in their focus on regular vs. irregular verb forms; the types of errors made are related to the learning path so defined.

1. INTRODUCTION

The first goal of the present study is to contribute to the descriptive study of Estonian morphosyntactic development. For this purpose we will present data from case studies of two boys, one (Andreas) a first-born child learning Estonian in a monolingual context in present day Estonia, the other (Raivo) a second-born bilingual child who learned Estonian in California over 25 years ago, exposed mainly to Estonian in the home but to English in day-care and the community at large.[1] We will note both similarities and differences in the course of early verb form learning in the two boys. By identifying individual learning paths we should be able to obtain a clearer picture of the effects of the particular ambient language structure on the range of possible approaches to first verb form learning in general.

The second goal is to extend to the acquisition of a language of a different morphological type questions of productivity in the early learning of verb morphology which have been raised in studies of Italian (Pizzuto and Casselli 1994) and Spanish (Gathercole, Sebastián and Soto 1999). For this purpose we will apply to our data the definitions of emergent contrast proposed by those studies and we will also consider the context of use of emergent contrastive forms to determine the extent to which evidence of context flexibility supports the evidence of contrast within morphological

N. Gagarina and I. Gülzow (eds.), The Acquisition of Verbs and their Grammar: The Effect of Particular Languages, 263–295.

paradigms in the first six months of verb use. We will then look at the occurrence of overgeneralization in spontaneous production and other errors of morphological usage, looking for indicators of generalized knowledge of the grammar or of emergent links between different verb stems and morphological forms. Finally, we will assess the relation between input frequency, in terms of both types and tokens, in one mother's input speech and the verb forms produced by her son.

In comparison with English, whose acquisition has been by far the most intensively studied to date, Estonian is a very highly inflected language. Although Estonian may be best known, along with both the closely related Finnic language Finnish and the distantly related Ugric language Hungarian, for its extensive nominal case system (14 cases: See Vihman 1982), the verbal system is no less rich in inflectional morphology (Erelt, Erelt and Ross 1997); neither system has received a great deal of attention from a developmental point of view (Dasinger 1997). Estonian also differs from the more inflected Indo-European languages, such as the Slavic languages or (in the case of the verb system) the Romance languages, for example, in that it is relatively agglutinative. Inflection is expressed exclusively by suffixation (and stress falls predictably on the first syllable in the core vocabulary (Tragel 2001); exceptions are relatively recent loan words, none of which are early learned verbs). Affixes typically express a single function each, i.e., mood, tense, person. There are no distinct verb classes or conjugations but merely stems and suffixes. Many stems are affected by internal changes known as 'consonant gradation' (with 'strong' and 'weak' grades), the phonological context for which is generally not transparent in Estonian, unlike Finnish, for example. The affixes are typically uniform, regardless of the stem to which they are attached, although many of the most commonly used core verbs show inflectional irregularities.

1.1. Adult verb system

Before embarking on our study of the children's development we briefly present the main features of the Estonian verbal system (see Appendix), focussing primarily on forms that appeared in the boys' production but providing an overview of the whole. The system includes several non-finite forms: two infinitives, traditionally known as 'first' (*-ma*) and 'second' (*-da*), and a set of four participles, present and past, active and passive.

The choice of the *ma-* vs. *da*-infinitive is determined by the control verb, e.g., *sa hakkad katsu<u>ma</u>* 'you're starting to try/touch', *sa hakkad söö<u>ma</u>* 'you're starting to eat' but *sa tahad katsu<u>da</u>* 'you want to try/touch', *sa tahad sü<u>üa</u>* [irreg. form of *söö*+*<u>da</u>*] 'you want to eat'.[2] The *ma*-infinitive is considered a supine, expressing 'an action following some other preceding action' (Erelt, Erelt and Ross 1997). Practically speaking, however, the

distinction between verbs requiring one vs. the other infinitive is somewhat arbitrary, as suggested by two of the formulations that children are likely to hear often: *sa pead katsuma* 'you have to try/touch', *sa pead sööma* 'you have to eat' vs. *sa ei tohi katsuda* 'you mustn't try/touch', *sa ei tohi süüa* 'you mustn't eat'. The infinitives are also subject to limited nominal inflection, but such uses are more common in narrative and formal language than in everyday colloquial discourse.[3]

The present active and passive participles are not part of the verbal system per se and were used by neither child within the period covered by this study. The past active participle is used in constructing perfect tenses, in combination with inflected forms of the verb 'to be' (as in Indo-European languages). It also serves as the (unchangeable) past tense form of the verb after the negative particle *ei* and as such is not uncommon in the input and is used to some extent by both children. The past passive participle is used in forming passive constructions; there is just one maternal and one child use in our data.

Among the finite forms there are four moods (indicative, imperative, conditional, evidential) and two tenses (present or 'non-past' and past); only the indicative has morphological tense marking. Note that there is no formal future tense; the future may be expressed by the present tense of verbs such as 'begin to' followed by the appropriate infinitive form, or simply by the use of future time adverbials and other expressions. There are three persons (but no grammatical gender, even in the pronoun) and two numbers. As in many Indo-European languages, the second person plural (p2) is used for polite address as well as for speaking to more than one person. In addition to the personal forms there is an 'impersonal', a single inflected form of the verb in both tenses.[4]

As in Finnish, the negative (indicative) is expressed by a negative particle followed by the verb stem. This particle, which is inflected for person in Finnish, has only the single form *ei* 'no, not' in Estonian. As a result, the present and past negative are unmarked for person, since person inflections are also missing from the verb stem. Note that negative replies in Estonian typically include the verb in one of its base (or 'stem') forms. Thus, a negative response to the question *Kas sa tule-d söö-ma* 'are you (sg.) coming to eat?' would typically be expressed as *ei tule* '(I'm) not coming', in which the pronoun is elided (despite the fact that person is nowhere overtly marked in the negative + verb sequence) but the main verb of the question is repeated in the negation. (Affirmative responses to yes/no questions also repeat the main verb of the question.)

In the imperative the second person forms are the only ones commonly used. The singular form, frequently used in speech addressed to a child, is a bare stem, distinguishable from the negative present tense indicative only by the co-occurrence of *ei* with the latter. In colloquial speech the first person

plural present tense (hereafter, 'p1pres') is generally used in place of the formal p1 imperative (ipv). In addition, there is a negative imperative particle inflected for all the forms of the imperative and followed by inflected imperative verb forms (see Appendix). Of these, only the unmarked s2ipv form (*ära* 'don't') occurred in the input speech we examined from one mother, and no identifiable negative imperative form was produced by either child.

Neither child made use of either the conditional or the evidential in the first six months of verb use; these forms are illustrated in an abbreviated manner in the Appendix.

2. METHOD

Our data are drawn from recording sessions supplemented by diary notes made by the authors. In Andreas' case 45-minute recordings were made on a roughly monthly basis from 1;8 to 1;11, followed by a period of daily one-hour recordings, 5 days a week, from 2;0.1 to 2;1.12 (see Vihman and Vija 2002). For Raivo, recordings were made at least once a month from 1;9 through 2;2, although the sessions vary considerably in length (from less than 15 minutes to over an hour). In order to assess the children's development on a comparable basis we include one or two recordings per month for each child (see Table 1).[5] The boys had produced only a single verb form each before the recordings began; this was the negative (*ei* + verb stem) in both cases, for one verb type (Raivo: *ei tee* 'no do', 'not do') or for two (Andreas *ei saa* 'can't', *ei taha* 'not want'). Thus the study covers the first six months of verb use for each child (as does Gülzow and Gagarina, this volume).

Table 1. Child age and MLU at recording sessions

	Andreas			*Raivo*		
Observation periods	*Age at session*	*Length of session (in minutes)*	*Mean length of utterance (in words)*	*Age at session*	*Length of session (in minutes)*	*Mean length of utterance (in words)*
1;8	1;8.10	45	1.2			
1;8.15	1;8.25	45	1.19			
1;9	1;9.11	45	1.5	1;9.10	19	1.08
1;9.15				1;9.20	19	1.18
1;10	1;10.3	45	1.56			

1;10.15	1;10.22	45	1.96	1;10.21	45	1.67
1;11				1;11.14	32	1.63
2;0	2;0.1	60	2.59	2;0.14	13	2.34
2;1	2;1.12	60	2.91	2;1.14	77	2.14
2;2				2;2.21	23	2.63

2.1. The identification of contrast (evidence of productivity)

In a study of Spanish verbal morphology Gathercole, Sebastián and Soto (1999) demonstrated the very gradual, 'piecemeal' development of verbal morphology by two children followed over a period of several months. The study focussed on identifying the first productive use of verbal inflection, applying for that purpose both separately (a more liberal version) and together (the more rigorous interpretation) the two-part criterion proposed by Pizzuto and Caselli (1994: 156) in a study of Italian verbal morphology:

(a) the same verb root appear[s] in at least two distinct forms, and
(b) the same inflection [is] used with at least two different verbs

For the present study we grouped the data into two-week periods, using the diary data to supplement the recordings. The earliest forms used by either child are unmarked stems, some of which are used appropriately, in the negative or 2d person imperative (s2ipv). Because our primary interest is to trace the emergence of the systematic use of inflections we distinguished between unmarked forms (negative, s2ipv) and overtly marked forms; we treated separately uses of the stem which occurred in lieu of a required inflected form (e.g., *kala uju* 'fish swim' for *kala uju-b* 'fish swims'). Following Pizzuto and Caselli (1994) and Gathercole, Sebastián and Soto (1999) we identified the emergence of contrast both for verb types (occurrence of more than one overtly marked form) and for inflectional categories (occurrence with more than one verb form). For this purpose we noted any relevant instances of an overtly inflected form (i.e., one or more tokens), in either diary or recordings. We also noted all the unmarked verb forms ('bare stems') produced by either child; these are not included in our tables of contrasting inflections but are discussed later.

3. RESULTS

3.1. Raivo

Although Raivo's first word combinations were recorded in the diary at 1;7.28 and his first verb form at 1;8.8 *ei tee* 'not do', he used few overtly marked forms in the first months of verb use. In fact, until 1;10 his only such form was the *da*-infinitive *saada* 'to get' (first recorded at 1;9.4). At 1;10 a second *da*-infinitive was used, but one which is morphophonemically opaque (irregular): *minna* 'to go'. (Most of the forms identified here as 'irregular' are the result of phonological assimilation across the stem/affix boundary, as in this case: *min(e)-* + *-da*.) Raivo's first contrasts by the first criterion – uses of a verb stem with more than one inflection – emerged only at 1;10.15, when he had made recorded use of 33 different verb types (see Table 2).

Table 2. Summary of verb types, inflected forms and emergent contrasts: Raivo

Child age	*1;9.0*	*1;9.15*	*1;10.0*	*1;10.15*	*1;11.0*	*1;11.15*	*2;0.1*	*2;0.15*	*2;1.0*	*2;2.15*
Verb types used	*5*	*12*	*16*	*27*	*21*	*17*	*18*	*19*	*32*	*23*
Cumulative verb types	5	16	21	33	41	46	47	51	57	60
Overtly inflected forms	*1*	*1*	*5*	*17*	*16*	*14*	*21*	*22*	*29*	*26*
Verbs used in 2+ forms (% of verbs)	0	0	0	5 (15)	5 (12)	6 (13)	9 (19)	10 (20)	16 (28)	18 (30)

Five common verbs, or 15% of Raivo's cumulative verb lexicon, show more than one inflected form in the two-week period beginning at 1;10.15. We will refer to this proportion as a 'contrast index', as it provides a way to relate verbs with contrasting forms to the cumulative lexicon to date. Not all of these verbs occur in the same inflected forms, however, and more importantly two of them involve irregular paradigms: the verb *ole-* 'to be' and the suppletive verb *mine-/lähe-* 'to go' (see Table 3). The third person singular present tense *on* 'is' (for the regular s3pres **ole-b*) is the only irregular verb form for that inflection; it is additionally unique in neutralising singular and plural number (i.e., *on* is also used in lieu of **ole-vad* '(they) are', the expected p3pres form). There is no evident (or phonetically transparent) relation between this form and either the other s3pres form that Raivo produced at 1;10, *sööb* 'eats', or the regular s2pres form *oled* 'you are'. The forms of the verb 'to go' derive from two distinct (suppletive) stems; the two infinitives and the forms of the imperative are all based on *mine-* while

the remaining forms use *lähe-*. Raivo's first two uses of the verb – *minna* 'to go' (*da*-inf) and *läheme* 'let's go, we're going' – have no surface features in common.

Table 3. Raivo's contrastive verbal paradigms

Contrastive verb uses are in plain font, stems and first uses in *italics*.

verb stems	*1;10.0*	*1;10.15*	*1;11.0*	*1;11.15*	*2;0.1*	*2;0.15*	*2;1.0*	*2;2.15*
and- 'give'						*annan*		annad
jää- 'stay'					jääda jäi			
käi- 'walk'		*käia*					käib	
leid- 'find'							leidsin	leidis
lõika- 'cut'			*lõikab*					lõikan
löö- 'hit'		*lõi*					löön	
mine-/lähe- 'go'	*minna*	läheme	lähevad	läks	lähen		läheb	
näge- 'see'				*näinud*			näen näed	
ole- 'be'	*on*	oled	olla olen	oli				
pane- 'put'		panna paned			pani	panen		
pida- 'have to'						*peab*	pean	
saa- 'get'	*saada*	sai			saad	saan		saab
söö- 'eat'	*sööb*				süüa söön sööme			
taht- 'want'			*tahan*			tahad		tahab
tege- 'do, make'	*teen*	teha tegi	teeb					teed
tule- 'come'		*tuli*		tulin		tulla		tuleb
vaata- 'look at, watch'						*vaatan*	vaatad	
võt- 'take'		*võtta*			võtan			

Taking the second criterion only, use of the same inflectional ending with more than one verb, we find just three inflectional forms used with more than one verb at 1;10.15 (Table 4): the *da*-infinitive (irregular: *käia* 'to walk, go', *panna* 'to put', *võtta* 'to take'…), s2pres *(oled* 'you are'*, paned* 'you put', *astud* 'you step'), and s3past *(sai* 's/he got'*, lõi* 's/he hit', *tegi* 's/he did, made'…). Of these forms only one, s2pres, is a regular form, with the transparently separable affix *-d*. Thus, it is not entirely clear to what extent productive knowledge should be imputed to the child even when both criteria are satisfied, as in the case of *on/oled, panna/paned, saada/sai, teha/tegi.* Of these, only *oled* and *paned* constitute the kind of regular, morphophonemically

transparent forms from which abstract knowledge of the morphological combinatory system could readily begin to be induced.

Table 4. Raivo's uses of verbal inflections

Second and later uses are in plain font, first uses in *italics***.**
For glosses see Table 3. Asterisks mark errors.

inflections	*1;10.0*	*1;10.15*	*1;11.0*	*1;11.15*	*2;0.1*	*2;0.15*	*2;1.0*	*2;2.15*
da-inf (irreg)	*minna*	käia panna võtta teha juua	olla		süüa tulla			
da-inf (reg)	*saada*			hoida peksta	jääda			
s1 pres	*teen*		olen tahan viskan		lähen mängin söön* võtan	annan panen saan vaatan ujun	ajan löön näen pean keeran näitan ostan otsin	lõikan
s2 pres		oled paned astud	tood viid		saad	tahad tõused	näed tahad vaatad	annad teed võid
s3 pres (irreg)	*on* 'is'							
s3 pres (reg)	*sööb*		lõikab teeb			peab	käib läheb nutab	kaevab saab tahab tuleb
p1 pres		*läheme*			sööme			
p3 pres			*lähevad*					
s1 past				tulin			leidsin	
s3 past (irreg)		sai lõi tegi tuli		läks oli	jäi pani			
s3 past (reg)			seisis ütles					leidis*
past act. part.		*kadunud*		näinud				

Looking at the next month, from 1;11 to 2;0, we find that only a single new verb (*tule*- 'come') is added to those showing contrast (Table 3). Since the number of new verbs produced continues to increase, the contrast index

decreases slightly, to 13%, at the mid-point of our longitudinal observations. As can be seen in Table 4, two new inflections now occur on multiple verbs, the remaining present tense singular forms, first and third. The second regular s3pres form occurs at 1;11.0, adding to the paradigm for *tege-* 'do, make', along with a new verb, *lõika-b* 'cuts', recorded in no other form until 2;2. *Ole-n*, s1pres of 'to be', also now occurs, contrasting with an earlier use of the same inflection marker for *tege-: tee-n* and two additional verbs, *taha-n* 'I want' and *viska-n* 'I throw'. Thus, by 2;0 five inflectional endings have been established as productive by the second criterion: the *da*-infinitive, all singular persons in the present tense and the third person of the past tense. Furthermore, both the *da*-infinitive and the third person past tense now have more than one regular as well as irregular exemplar (*hoi-da* 'to hold', *peks-ta* 'to spank' in addition to *saa-da*, and *seisi-s* 'stood', *ütle-s* 'said').

Remarkably, although increasing numbers of verbs occur in one or another of the established inflections in each successive period (Raivo's contrast index is 20% by 2;0.15, 30% by 2;2), very few new inflectional markers are established as productive. Only two verbs are inflected for p1pres, *söö-me* 'we're eating/let's eat' and *lähe-me* 'we're going/let's go'. The verb *leid-* 'find' occurs in both s1past and s2past, *leid-s-in*, *leid-s-id*; each of those person forms occurs in a single additional (irregular) verb, *tuli-n* 'I came' and *pan-i-d* 'you put'. Only one more distinct verb inflection is used, the past active participle *-nud.* It occurs just once as a past tense negative verb stem *ei näi-nud* 'didn't see'; an earlier use is the purely adjectival *kadu-nud* 'lost', i.e., 'all gone'. Thus, by the end of six months of verb use Raivo has used 60 different verb types, almost a third of them in more than one form, but only eight distinct inflectional forms have been used for more than one of these verbs, and only seven verbs occur in more than three different forms (the earliest occurring verbs with contrast, *mine-* 'go', *ole-* 'be', *pane-* 'put', *saa-* 'get, become', *tege-* 'do, make', and also *söö-* 'eat' and *tule-* 'come').

3.2. Andreas

Andreas began verb use about a month earlier than Raivo, and in the six-month period of this study he produced over 100 different verb types and 89 different overtly inflected forms (see Table 5). His more prolific output should thus provide us with a good opportunity to see whether a child growing up as a monolingual Estonian speaker, amply exposed to his native language outside as well as inside the home, will show more rapid or more dramatic ('across the board') learning of the complex set of verbal inflections.

Table 5. Summary of verb types, inflected forms and emergent contrasts: Andreas

Child age	*1;8*	*1;8.15*	*1;9*	*1;9.15*	*1;10*	*1;10.15*	*1;11*	*1;11.15*	*2;0*	*2;1*
Verb types used	*11*	*10*	*36*	*19*	*42*	*38*	*11*	*9*	*54*	*46*
Cumulative verb types	11	18	45	52	70	78	83	86	107	113
Overtly inflected forms	*0*	*1*	*13*	*12*	*31*	*40*	*13*	*10*	*73*	*89*
Verbs used in 2+ forms (% of verbs)	0	0	1 (2)	3 (6)	9 (13)	18 (23)	20 (24)	23 (27)	29 (27)	41 (36)

Like Raivo, in his first few weeks of verb use Andreas produced mainly verb stems, some of them identifiable as imperatives (*loe* 'read!') or negatives (*ei oska* 'don't know how', *ei saa* 'can't', *ei taha* 'not want'). By the end of the two-week period beginning at age 1;9.0 he had produced 45 different verb types but had used more than one overtly inflected form for only a single verb, *tuli* 'came': *tuleme* 'we're coming', and in the following fortnight only two more contrastive forms were added (*sai* 'got': *saab* 'gets'; *tõi* 'brought': *toob* 'brings'), giving a contrast index of 6%. Between the ages of 1;10-1;10.15, the mid-point of the study, Andreas produced nine verbs in more than one inflectional form, out of 70 different verb types produced (13%: see Table 6) – the same proportionate level of contrast identified at Raivo's mid-point (Table 2).

Table 6. Andreas' contrastive verbal paradigms

Contrastive verb uses are in plain font, stems and first uses in *italics*.

verb stems	*1;8.15*	*1;9.0*	*1;9.15*	*1;10.0*	*1;10.15*	*1;11.0*	*1;11.15*	*2;0.0*	*2;1.0*
aita- 'help'		*aitas*			aitab				
and- 'give'									annab andis
ehita- 'build'									ehitada ehitab ehitame
hakka- 'begin'					*hakkab*			hakkavad	
helista- 'call'								helistab helistas	
hoid- 'hold'				hoiab hoidis					hoida
istu- 'sit'				*istus*	istuma istub			istud istuda	
joo- 'drink'				jõi joob juua					
joonista- 'draw'					*joonistab*			joonistame	
kadu- 'be lost'		*kadus*			kadusid				
keera- 'turn (tr.)'				*keeras*					keerab
kirjuta- 'write'					*kirjutama*		kirjutab		kirjutatud
kisku- 'touch'							kiskuda kisub		
kopsi- 'knock'			kopsis		kopsida			kopsib	
kukku- 'fall'		*kukkus*			kukkusid				kukkuda
käi- 'walk'				*käisin*	käisid		käis		
laul- 'sing'					*laulis*			laulavad laulda	
leid- 'find'					*leidis*				leiab
luge- 'count, read'						loeb		lugema	lugenud
						luges			
lõika- 'cut'				*lõikama*				lõikas	
maitse- 'taste'									maitseb maitses
mine/lähe- 'go'		*läheme*		läks	minna			lähed	
								lähevad	
mängi- 'play'							mängima mängis	mängib	mängida mängid

nuusuta- 'sniff'							nuusutab	
							nuusutada	
näita- 'show'							näitama	näitab
								näitas
ole- 'be'		*on*						olla
								olema
								oli
oota- 'wait'						ootas		ootab
oska- 'know how'					*oskab*			osata
otsi- 'look for'					*otsib*			otsima
pane- 'put'		*pani*		panna				panema
								paneb
								paneme
pese- 'wash'		*pesta*		pesi			peseb	pesevad
							pesema	
põle- 'burn'				*põleb*	põlema		põlevad	
saa- 'get'		*sai*	saab				saada	saad
								saanud
söö- 'eat'				sööb			süüa	sõi
				sööma				
taht- 'want'					*tahab*			tahtnud
tege- 'do, make'				*tegi*	teeb		teha	teed
					teeme			teinud
					teevad			
					tegema			
too- 'bring'		*tõi*	toob					toon
tõst- 'lift'				*tõstis*				tõstab
tule- 'come'	*tuli*	tuleme		tulge	tuleb	tulevad	tulla	
					tulnud			
vaata- 'look at, watch'						vaatab		vaadata
						vaatas		
võt- 'take'					võtad		võtma	võtavad
				võtab			võtta	võtame

It can be seen from Table 7 that Andreas concentrated his earliest uses of marked verb morphology on a smaller number of different inflections than Raivo did. The irregular third person past tense form, recognizably marked with *-i,* is one of the contrasting forms in all three of the first verbs Andreas produced in more than one inflected form and the remaining forms are all regular, s3pres (*-b*) and p1pres (*-me*). The pattern of use of a small number of inflections on increasing numbers of verbs continues, with third person

occurring in both present and past tenses for new verbs in each session from 1;9.15 on, as is particularly evident from the periods in which recordings were made. Another form comes to be of common occurrence from 1;10 on, namely, the *ma*-infinitive, which failed to occur at all in Raivo's first six months of verb use. In comparison with Raivo's verb forms the relative frequency of the irregular vs. the regular forms of both the *da*-infinitive and s3 past is reversed: Andreas has a combined total of 19 irregular forms and 52 regular forms for these two inflections – vs. Raivo's 17 irregular and 7 regular forms. This means that the regular forms, from which the grammatical ending can more readily be extracted for morphophonemic reasons, are well practiced in Andreas' case, given the steady blossoming of s3 present and past and *ma*-infinitive forms. However, no other forms occur for more than two or three verbs each until the last two months of the study when four new verbs each occur in s2pres and p3pres and five occur with the past active participle. Interestingly, only two uses of the first person were recorded in all, one each in the past tense (*käi-si-n* 'I walked', at 1;10) and the present (*too-n* 'I'm bringing', at 2;1). By the end of the period of observation Andreas had produced 113 different verb types, almost twice as many as Raivo, yet his contrast index (36%) remains close to Raivo's (30%).

Table 7. Andreas' uses of verbal inflections

Second and later uses are in plain font, first uses in *italics*. For glosses see Table 6. Asterisks mark errors.

inflections	*1;8.15*	*1;9.0*	*1;9.15*	*1;10.0*	*1;10.15*	*1;11.0*	*1;11.15*	*2;0.0*	*2;1.0*
da-inf (irreg)				juua	minna			süüa	olla
				panna				teha*	
								tulla	
								võtta	
da-inf (reg)		pesta			kopsida		kiskuda	istuda	ehitada
		kiikuda						laulda	hoida
								nuusutada	kukkuda
								pühkida	mängida
								remontida	osata*
								saada	vaadata
								visata*	
								värvida	
ma-inf				lõikama	istuma		mängima	lugema*	kuulama
				sööma	kirjutama			näitama	lükkama
				lõikama	põlema			pesema	olema
					tegema			valvama	otsima
					rääkima			võtma	panema
					sõitma				
p2ipv				*tulge*					

s1 pres									toon
s2 pres					*võtad*			istud	mängid
								lähed	saad
									teed
									võimled
s3 pres (irreg)		*on*							
s3 pres (reg)			saab	hoiab	aitab	loeb	kirjutab	helistab	annab
			toob	joob	hakkab	vaatab	kisub*	kõnnib	ehitab
				nutab	istub		lehvitab	kopsib	jookseb
				põleb	joonistab			küpsetab	kardab
				sööb	koristab			kuuleb	keerab
				tudub	oskab			mängib	kraadib
				tõuseb	otsib			nuusutab	laseb
				võtab	tahab			parandab	leiab
					teeb			peseb	maitseb
					tuleb				näitab
									ootab
									paneb
									tõstab
									tudub
p1 pres		läheme			teeme			joonista-me	ehitame
		tuleme							kraadime
									paneme
									võtame
p3 pres					*teevad*			hakkavad	pesevad
								lähevad	saavad
								laulavad	võtavad
								põlevad	
s1 past				*käisin*					
s2 past					kadusid				
					kukkusid				
					käisid				
s3 past (irreg)	tuli	pani		läks				jäi	oli
		sai		jõi					sõi
		tõi		pesi*					
				tegi					
s3 past (reg)		aitas	ärkas	hoidis*	laulis	luges*	käis	helistas	andis
		kadus	köhis	istus	leidis	ootas	mängis	kinkis	maitses
		kukkus	kopsis	keeras	lendas	sadas*		küünistas	näitas
		ostis	lõppes	tõstis*	viis*	vaatas		lõikas	
				korjas				tilkus	
				kutsus				väsis	
				väsis					
active part					*tulnud*				lugenud

		saanud
		tahtnud
		trüginud
		teinud
passive		kirjutatud
impersonal	*küsitakse*	

Figure 1 provides a comparison of the rate of cumulative verb learning shown by each child with the rate of increase in the contrast index.

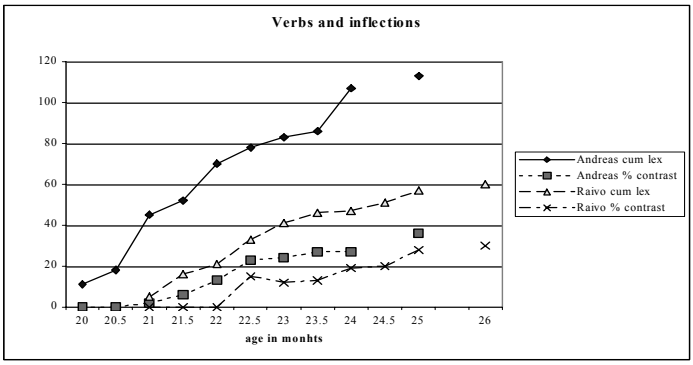


Figure 1. Rate of verb learning vs. rate of emergence of morphological contrast.

'Cum lex' = cumulative lexicon; % contrast is the number of verb types used with more than one inflectional form.

It is apparent that while Raivo's Estonian verb learning is much slower, the two boys differ very little in their rate of learning of contrast.[6] This suggests that a learning process must be involved that is the result of neither the automatic unfolding or 'maturation' of an underlying blueprint for grammar nor the simple accumulation of lexical knowledge. On the one hand, no sharp increase in inflected forms is in evidence to reflect the 'triggering' of functional knowledge which is predicted by the maturational account. On the other hand, it is not the case either that contrast emerges in direct parallel with verb learning but with a time lag sufficient to allow pairs of verb stems or inflections to be produced and identified. Instead, the gradual, 'piecemeal' learning of verbs appears to be followed by a process of abstraction or induction of the underlying structure, resulting in a separate, even more gradual learning curve for morphological contrast for both children, with

some months of experience with production required before contrast is seen. This supports the idea of grammatical knowledge having to be abstracted out of language use.

3.3. Contrast in contextual use

It is possible to look at the emergence of contrast in actual use in context, although this is difficult to quantify or to operationalise as an objective method. Nevertheless, we will draw here on Andreas' more copious material to see to what extent contrastive usage can be observed as it emerges, and to what extent such usage agrees with our findings based on analyses of multiple inflections per verb type. One of Andreas' earliest within-paradigm contrasts is s3pres vs. s3past of *too-* 'bring'. The past form was used already at 1;9, e.g., *Mana* [for *vanaisa*] *tõi kuusepuu* 'Grandfather brought the Christmas tree' (s3past), the present at 1;9.15, e.g., after fetching a fruit from the kitchen, *Antsu toob* 'Antsu brings/is bringing (it)' (s3pres). Similarly, the verb *saa-* 'get, become' is used first in the past at 1;9 (*otsa sai* 'came to an end', lit. '*got* into-end' (s3past), a collocation equivalent to 'allgone' in English), but occurs in the present at 1;9.15: *kätte saab* 'gets it, catches it' (lit., 'gets into-hand' (s3pres), reaching for something on a high shelf). The context for these verb uses in the present vs. the past is rather different, both of them possibly situationally primed and so providing only weak evidence of understanding of the distinct forms' potential for contrastive use.

Some uses are even more tightly associated with particular routines, such as Andreas' use of the p2ipv for only a single verb, *tule-* 'come', as in *tulge sööma!* 'come (pl.) eat!' (addressed to his father as mother is cooking, at 1;9: note that the use is inappropriate for a single interlocutor, but reflects the mother's usage, addressing both the child and his father; it is a formulaic invitation to a meal), *tulge sisse* 'come (pl.) in', addressed to anyone requesting entry at the front door, and the attempted paraphrase, at 1;10:

(1) M: *Tulge* nüüd pannkooke sööma!
'*Come* (pl.) eat pancakes now!'
A [to his father]:
Kutsus emme *tulge* panni sööma
call-s3past mummy come-p2ipv pan[cake]-sPart eat-ma-inf
'Mother called, *come* eat pancakes'

When, at 1;10.15, the child uses the contrasting s3pres form *issi tuleb* 'daddy's coming' as he hears a key turn in the lock, the use is again closely tied to the context. Already at 1;10, however, we find appropriate contrastive use of the verb *joo-* 'drink' in three different forms in three utterances

produced in sequence, appropriately referring differently to a single event (Andreas uses the nickname Antsu to refer to himself):

(2) Antsu *juua* veel
Antsu drink-da-inf more
'Antsu [wants] *to-drink* some more' (requesting a drink from his mother, in the kitchen, using the *da*-infinitive which would be appropriate after the missing verb *taha-* 'want')

(3) Antsu *joob*
Antsu drink-s3pres
'Antsu *is drinking*' (s3pres, addressing his mother in the kitchen)

(4) Antsu *jõi* juba
Antsu drink-s3past (irreg.) already
'Antsu already *drank*' (in the next room, addressing his father)

Some additional examples of appropriately contrastive inflectional choices in closely related situational contexts:

(5) küünal põleb [1;10]
candle burn-s3pres
'the candle *is burning*' (after turning the switch on an electric candle)

(6) suur *põlema* [1;10.15]
big (one) burn-ma-inf
'[set] big [one] *to-burning*' (entering the dark bedroom, requesting light, using the *ma*-infinitive which would be appropriate after the missing verb *pane-* 'put, set')

(7) Antsu *aitas* [1;9]
Antsu help-s3past
'Antsu *helped*' (after energetically participating in clean up with his mother)

(8) Atsu *aitab* [1;10.15]
Antsu help-s3pres
'A(n)tsu *is helping*' (gives mother a book, looking for a particular picture)

In the final recording session at 2;1 we find immediate contrastive use of s3pres and s3past for the verb *maitse-* 'taste', not previously recorded:

(9) tita, tita saab mampsu [*for* ampsu]
dolly, dolly get-s3pres bite
'dolly, dolly gets a bite' (as he gives play food to the doll)

Atsu ka *maitseb*
Antsu also taste-s3pres
'Antsu *is tasting/will taste* (it) too'

tita ka *maitses* ära juustu
dolly too taste-s3past up [perf. marker] cheese
'dolly *tasted* the cheese too'

Note that without experimental evidence (using nonsense forms, for example) there is no way to draw a line indicating just when a given inflection begins to be used productively, but these examples of flexible child use in situations which allow for some choice of possible forms provide a suggestive illustration of the move from contextually tied usage to what appears to be more flexible, probable contrastive knowledge.

3.4. Errors of omission and commission, form and function

Three major types of child production errors may be distinguished in the acquisition of Estonian verbal morphology. First, there are errors of omission: the use of the verb stem alone, i.e., of an unmarked form of the verb, where an inflection is clearly required by the intended meaning, based on context. Second, there are overt errors of form (errors of commission), where either the stem or the inflectional suffix is different from the target form, usually suggesting phonological or morphological overgeneralization. Finally, there are overt errors of usage or function (errors of commission again), where one inflected form is used in place of another. We take up these errors in turn.

3.4.1. Use of stems in lieu of inflectional affixes

As noted earlier, the bare verb stem, which occurs in both the second person singular imperative and the negative present tense in all persons, serves as the default verb form for both children. As verb forms begin to be learned, the use of the stem decreases, although it alternates with inflected forms throughout the period of this study in Raivo's case and for much of the period in Andreas' case.

Even at 2;0, when he has produced 46 different verb types and is using a mean of 15 different inflectional affixes in each two-week period (with 19% of his verbs having been produced in at least two contrasting inflections), Raivo substitutes the stem for a range of different target suffixes: all forms of the present tense singular, the first and third person plural, the third person past tense, and both infinitives. The range of verb forms which Raivo produces in stem form only is not unrestricted, however: After 2;0 all of the target forms produced as stems are *regular da*-infinitive and past tense forms. In the following examples the target form is retrievable from the context; note the cooccurrence of correctly inflected finite and non-finite forms of some verbs alongside uses of the stem only for others:

(10) Olen maas *vaata* juttu [2;0.12]
be-s1pres ground-inessive look at story
'I'll be on the floor, [I'll] *look at* a story' (for *vaata-n*: getting down off couch. Raivo uses s1pres *olen* 'I am' correctly here; the context provides the future meaning glossed in English)

(11) mul on vaja *pese* kätt [2;0.19]
me-sAdess be-s3pres necessary wash hand-sPart
'I need [to] wash my hands' (for *pes-ta*)

(12) ma tahan minna Virve voodi [peale] *maga* [2;0.21]
I want-s1pres go-da-inf Virve bed [onto] sleep
'I want to go [onto] Virve's bed *sleep*' (for *maga-ma*)

(13) Brendan *kukku* eile maha [2;0.21]
Brendan fall yesterday down
'Brendan *fall* down yesterday' (for *kukku-s*)

These uses of unmarked stems for verbs with regular endings constitute striking evidence of Raivo's *lack* of system learning to this point. By 2;0 the child had produced several verbs in each of the required inflectional forms, so that by liberal criteria he could be taken to 'know' the inflectional endings. The evidence suggests that the knowledge is still being acquired verb-by-verb, however.[7]

Andreas shows a leap in production of inflectional markers between 1;10, when he has produced 70 different verb types (contrast index 13%) and up to 31 different inflectional affixes in a two-week period, and 1;10.15, when he adds just eight new verb types (for a total of 78) but produces 40 inflected forms (contrast index 23%). Out of a total of 39 verb forms used, 13 (.33) still occur in stem form rather than the required inflection. In comparison, at 2;0,

when Andreas has produced 107 verb types (contrast index 27%), just 3 of the 54 verb types are stem forms instead of the target form. A review of the specific uses Andreas made of the stem during the recording at 1;10.22 shows that all but one of the misuses were in reference to himself, sometimes in alternation with s3 in the same discourse, or to others (his mother or father, the cat); all uses of stem for inflected form involve a target present tense form. In later sessions use of the uninflected stem for self is replaced by consistent use of s3pres; it is only in the last recording used here, at 2;1.12, that Andreas produces his first s1pres (see Table 7). Examples of his use of bare stems for self-reference at 1;10.22:

(14) Antsu *istu* maha 'Antsu *sit* down'
Antsu sit down
(*for* istu-n [s1pres])

(15) valge kiisu *tudu* 'White kitty *sleep*'
white kitty sleep (baby-talk verb)
(*for* tudu-b [s3pres])

(16) M: mis sa teed nüüd? 'What are you doing now?'
A: *tants* 'dance'
(*for tantsi-n* [s1pres])

3.4.2. Phonological errors

The consonant changes ('consonant gradation') found in stems and, to some extent, in affixes in the adult language are not phonologically transparent. (Vowel changes also sometimes occur.) In particular, the 'weak grade' *-d* of the da-infinitive may assimilate to the stem (as in *min-na* 'to-go' for **min-da, tul-la* 'to-come' for **tul-da*); it may also occur in the 'strong grade', as *-ta:* cf. *vasta-ta* 'to answer', *osa-ta* 'to know how'. In contrast, the *ma*-infinitive is always regular. Our tables showing all the target forms attempted by either child are intended to help the non-Estonian reader to appreciate the extent of phonological alternation in the verb morphology.

It is striking to observe that Raivo seldom made phonological – or more accurately, morphophonological – errors in this period, whereas he made large numbers of morphological substitutions, as described below. Only three morphophonological errors could be identified, two from recordings:

(17) ma *süüan* kaelkirjak [2;0.21]
I eat-s1pres giraffe
'I *eat* giraffe' (has plastic giraffe in his mouth)
(for *söö-n*: Uses *da*-inf form *süüa* as if it were the verb stem, adding the first person marker *-n*; he also omits the partitive case marker from the noun)

(18) mina *luka* [jüka] üks kaks [2;1.14]
I count one two
'I *count* one two'
(for *loen*, using an approximation of the strong form of the stem *luge-* here, appropriate for the two infinitives and the past but not the present, imperative, or negative forms of the verb)

(19) Vilve *leidsi* seda [2;2.21]
Virve find-s3past this-sPart
'Virve found this'
(for *leidis*; this is a back formation from the s1past and s2past: *leidsin, leidsid*)

In short, Raivo was observed to make few such errors, but those few occurred in the last months of the study and show the effects of different forms of the verb on the intended form. This shows that connections are being made at this point; the errors reflect the emergence of system, like the classic examples of regression in the learning of English past tense.

Given the evidence already provided of Andreas' greater progress toward a productive inflectional system within the period of the study we expect to see more phonological errors, and we expect to see them in the last half of the study. This is just what we find. The first errors were recorded at 1;10, both in the diary and in transcriptions of recorded sessions. All those errors involve s3past. In most cases we see overgeneralization of the present tense stems: **ostas* 'bought', **hoias* 'held', **loes* 'read', **sajas* 'precipitated (rain, snow)' and **tõstas* 'lifted' for *ostis, hoidis, luges, sadas* and *tõstis.* In addition, at 1;10.6 Andreas produces a regularized past tense form **pesis* '(s/he) washed' for the irregular *pesi* and at 1;10.17 the overmarked disyllabic **viisis* '(s/he) took away' for the regular *vii-s*, adapting the unusual monosyllabic past tense form to fit the shape of all of his other regular verbs. Finally, at 1;10.22 he produces twice in succession the same irregular form arrived at by Raivo, **leidsi* for *leidis,* immediately following his mother's use of the s2pres form on which the overregularisation is modelled:

(20) M: näe, leidsid ülesse!
'See, you found it!' (after A locates a cork under a chair)
A: *leidsi* ise # *leidsi* ise Atsu
find-s3past self find-s3past self Antsu
'*Found* it [my]self. Antsu *found* it self'

In the following month errors affect other verb forms: The irregular s3pres *on* 'is' is regularized (**oleb*), a target 'weak grade' present tense *kisub* is formed on a 'strong grade' stem instead (**kiskub* 'meddles with, touches'), and alternating stems are interchanged in producing the irregular verbs *mine-/lähe-* 'to go' (suppletive) and *tege-/tee-* 'to do, make'. From 1;11.23 on we find errors of overregularization in the *da*-infinitive, only rarely produced in the preceding months (**hoiada* 'to hold' for *hoida,* **lauldada* 'to sing' for *laulda,* **oskada* 'to know how' for *osata,* **tegeda* 'to do, make' for *teha* and **viskada* 'to throw' for *visata*), and one mischoice of stem for a *ma*-infinitive, the weak grade being generalized from the present tense: **loema* 'to read' for *lugema*.

3.4.3. Morphological errors

In Raivo's data we find a large number of cases of 'functional errors', in which the morphological form produced is wrong for the intended meaning. These include both within-paradigm substitutions, such as uses of the wrong person within a tense, and across-paradigm substitutions, including uses of the infinitive for a finite form or of affirmative for negative stems and vice versa. The errors began to be seen only shortly before the first phonological error was recorded, at 1;11.26: *Virve tahan* 'Virve want' (lit. 'I-want') (instead of the s3pres *tahab,* obligatory after mention of a third-person subject, his sister Virve: Raivo was trying to relay a request of his sister's); the same error occurred again a few days later. More surprising is his use of p1pres *lähe-me* 'let's go, we'll go' for s3pres *lähe-b* 'goes' and also in place of the *da*-infinitive *min-na*:

(21) ei taha putukas *läheme* sisse [2;0.1]
no want insect go-p1pres to-in
'don't want insect *let's go* in', i.e., 'We don't want insects to come in' (noting the fact that the door to the porch is open) (for *läheb,* likely influenced by the formulaic expression, *läheme sisse* 'let's go in')

(22) a tahan *läheme* kooli [2;0.7]
[I?] want-s1pres go-p1pres school-sIll
'I want *let's go* to school' (for *minna*)

(23) Virve *läheme* oma kooli [2;0.26]
Virve go-p1pres own school-sIll
'Virve *let's go* to her school', i.e., 'Virve is going to school'

From 2;0.19 on we find use of the third person alternating with the first person to refer to himself. The remaining errors involve uses of known forms where a heretofore unused form is required: *Sina *on paha poiss* 'you is a bad boy', using *on* (s3pres) where *oled* (s2pres) is required; *ema peab kätte *saa-da* 'Mother must to-get it', using the *da*-inf where the *ma*-inf is required. Finally, at 2;1.14 we have the use of s3past where p3past is required by the plural subject, *Need *tuli OUT.* 'They came OUT'.[8]

Raivo's across-paradigm substitutions occurred in the same time period, beginning with a use of the affirmative past tense for a negative: 1;11.22 Raivo ei **läks* OUTSIDE 'Raivo not *went* OUTSIDE': The required form is the perfect active participle *läinud,* so far produced only for the verb *näge-* 'see': *näinud.* Although he was currently producing this form of that verb for the past negative, it appears that he had not yet made the across-verb connection at this point. On the other hand, at 2;0.21 we find the *da*-infinitive being used in lieu of the same s3past form: BALL **minna* vette 'BALL *to-go* in the water'. Similarly, at 2;0 we find two cases of the affirmative s1pres for the negative, where the correct form was used two months earlier: ei **lähen* vanni 'I not *go* to bath' (i.e., not take a bath). In another attempt at a past tense negative he made use of the *da*-infinitive:

(24) üks *jäi* kooli [2;0.13]
one stay-s3past school-sIll
'[lit.] one *stayed* at school', i.e. 'got left at school' (commenting on seeing only one shoe)
üks ei *jääda* kooli
one no stay-da-inf school-sIll
'one not *to-stay* at school' (for *jää-nud*)

The large number of errors made in this period argue against rote learning of whole contextually tied utterances in Raivo's case; the same point has been made in relation to this child's heavy mixing of English words into his Estonian discourse in his first combinations (Vihman 1999). On the other hand, the data do provide further evidence for an extended period of whole-

word usage in lieu of analysis into stem + affix (for evidence based largely on nominal forms, see Vihman 1982). In several of the instances noted here Raivo seems to have been using, for any one period (days, weeks), just one or two (holistic) forms of each verb. Whereas some uses may directly reflect his experience of adult usage (as with (24) 'got left at school'), other uses show a surprising lack of sensitivity to the requirements of particular constructions or contexts (e.g., the past tense form to be used after the negative particle *ei*).

Creative work on Raivo's part – influenced by the adult system but never directly experienced as part of the input – can be seen in his invention of a non-existent affirmative form, constructed on analogy with the negative, presumably, with the word *jaa* 'yes' preposed to the verb form as the word *ei* 'no' is preposed for the expression of the negative[9]:

(25) ei pime vaja, *jaa* pime vaja [1;10.23]
no dark needed, yes dark needed
'no dark needed, *yes* dark needed' (turning lights off and on)

(26) Virve *jaa* tahan dussi alla [2;0.3]
Virve yes want-s1pres shower to-under
'Virve *yes* I-want [to go] into the shower': contrasting his sister's preferences with his own) (for s3pres *taha-b*; cf. (23))

(27) see ei kuku, kuku maha … see *jaa* kuku maha [2;1.14]
this no fall, fall down this yes fall down
'this not fall, fall down, (just a few minutes later) 'this *yes* fall down' (stem for s3pres)

(28) mul on *jaa* tasku [2;1.14]
me-sAdess be-s3pres yes pocket
'I have *yes* pocket' (i.e., 'I do too have a pocket')

(29) see ei tööta… see *jaa* tööta [2;2.22]
this no work… this yes work
'this doesn't work' (disappointed in the crayon he's using);
(just a few minutes later) 'this *yes* work' (stem for s3pres)

Andreas' morphological errors are of a far more restricted type, most of them reflecting his confusion over first-person reference, the kind of error better known from pronominal 'I/you' confusion in English. Beginning at 1;10.15 we have several uses of s2past for s1past (*kukku-si-d* 'you fell' for *kukku-si-n* 'I fell', *käi-si-d* 'you walked' for *käi-si-n* 'I walked') or s2pres for s1pres (*istu-d* 'you're sitting' for *istu-n* 'I'm sitting', *lõika-d* 'you're cutting'

for *lõika-n* 'I'm cutting', *ole-d* 'you are' for *ole-n* 'I am', *otsi-d* 'you're looking for' for *otsi-n* 'I'm looking for'). At the same time he more typically refers to himself in the third person, frequently including his name. Even at the end of this study, when Andreas' MLUw of 2.6 reflects his use of quite long and complex sentences, he seldom makes use of pronouns. Of the few remaining errors two involve agreement:

(30) külalised *käisime* [2;0.1]
guest-pNom walk/come-p1past
'lit. guests we-came', intending 'guests came' (for p3past *käi-si-d*) (perhaps influenced by the formulaic *külas käi-si-me* 'we went visiting', lit. 'we walked in the village')[10]

(31) ilus lamp *põlevad*
nice lamp burn/shine-p1pres
'nice lamp *are shining*' (for s3pres *põle-b*)

In addition we find one case of an affirmative s3pres being used after the negative: 1;10.20 *ei joonistab* 'not s/he-draws', for *ei joonista*, and a use of the *ma*-infinitive where the *da*-infinitive is required, after *tahab* 'wants'.

3.4.4. Verb types and tokens in input speech

Since only Andreas' mother is a native speaker of Estonian we confine our analysis of adult input to her speech. For this purpose we analyzed seven tape recordings (45 minutes each) made in the period before intensive weekly recordings began, when Andreas was aged 1;8 to 2;0. Table 8 indicates the distribution of inflectional forms in both types and tokens, including the verb stem as used for the expression of either s2ipv or present tense negative.

Table 8. Distribution of verb inflections in Andreas' mother's speech (ranked in order of token frequency)

Inflectional form	*types (prop.)*	*tokens (prop.)*	*Inflectional form*	*types (prop.)*	*tokens (prop.)*
stem alone	70 (.17)	604 (.26)	past active participle	8 (.02)	14
s3pres/irreg.	1	569 (.24)	s1past	4	7
s3pres/regular	79 (.19)	333 (.14)	p2past	4	11
s2pres	48 (.12)	206 (.09)	p1past	8	9
p1pres	29 (.07)	114 (.05)	-mas	3	7
-da infinitive	26 (.06)	89 (.04)	p3past	4	6
s3past/irreg.	8 (.02)	83 (.04)	p2pres	1	3
s3past/regular	29 (.07)	71 (.04)	present impersonal	2	3
s2past	20 (.05)	69 (.03)	past impersonal	1	2
p3pres	25 (.06)	58 (.02)	past passive participle	1	1
ma-infinitive	18 (.04)	39 (.02)	conditional	1	1
s1pres	21 (.05)	33 (.01)			
Total				411	2333

If we group regular and irregular verb forms together the two highest frequency forms are the same for both types and tokens: The third person present tense accounts for 39% of all verb uses in tokens, 19% in types, while the stem alone accounts for 26% (tokens) or 17% (types). Division of third person verb forms into regular vs. irregular, for both indicative tenses, provides a more useful account of the potential impact of input frequency, however, since the relationship between regular and irregular verb forms exists only at an abstract grammatical level: As noted earlier, the sole irregular s3pres form (*on* 'is/are', the only present tense form with syncretism of number) bears no phonological relationship to the other s3pres forms, which take the ending -*b*. Looking at Andreas' earliest overt markers, we see that the form *on* 'is, are', which alone accounts for 24% of his mother's verb uses (tokens), occurs in his first month of production of marked inflections, at 1;9, although no other forms of the verb *ole-* 'to be' occur until the last period of the study, at 2;1. This appears to be a simple frequency effect.

The next most frequent form in the input is the regular third person present tense (14% tokens, 19% types), which Andreas first produces at 1;9.15 and which contrasts with the past tense forms of some verbs at 1;10. However, at 1;9 Andreas already produces the regular past tense of several verbs (4% tokens, 7% types in his mother's speech), and he also produces *ma*-infinitive forms at 1;10 (only 2% tokens, 4% types in the input). Recall too that the *da*-infinitive, which is of slightly higher incidence in his mother's

speech (4% tokens, 6% types), occurs in Andreas' production for a smaller number of verbs than the *ma*-infinitive over the course of the study. Similarly, Andreas made almost as much use of the past active participle as of the p1pres over the study as a whole, although his first p1pres forms were produced already at 1;9, his first perfect form only at 1;10.15. In the input, p1pres accounts for 5% of all verb tokens while the past active participle accounts for less than 1%. Only in the case of a morpheme of exceptionally high frequency – the s3pres of 'be' (in contrast with other forms of the copula) – does input frequency clearly seem to have played a direct role.

4. DISCUSSION AND CONCLUSION

We have traced the emergence of the first verbal inflectional markers in two children acquiring Estonian. In spite of the fact that the learning conditions were different in two important respects – monolingual vs. bilingual exposure and first- vs. second-born child – we found a number of similarities. Both children began with unmarked verb stems, which are modelled in the input in two frequently occurring grammatical forms, the second person singular imperative and the present tense negative. Additionally, both children showed a very gradual increase in overt inflectional forms, with different verb types sometimes being used for a period of months with different individual inflections, and individual inflections similarly occurring on no more than one or two verb types, rather than morphemes being produced for several verbs at a time soon after their first use. This finding is in agreement with the earlier work of both Pizzuto and Caselli (1994) and Gathercole, Sebastián and Soto (1999) on the acquisition of Italian and Spanish verb morphology, respectively.

By charting growth in the cumulative lexicon alongside the emergence of within-paradigm contrast we were able to see that despite the fact that Andreas learned almost twice as many verb types in his first six months of verb use, the rate of learning of contrast in the two boys was quite closely matched. Furthermore, by looking at Andreas' uses of verb forms in context we were able to validate, to some extent, the usefulness of the proposed criteria for identifying contrast. Despite the fact that some verbs did occur in more than one inflectional form in the early months of the study, no convincing evidence of use of any verbs in genuinely contrastive contexts could be found until 1;10.15, when Andreas' contrast index had reached .23 (and his MLUw almost 2.0). After that point several instances of contrastive use (present vs. past, finite vs. non-finite verb, singular vs. plural third person) could be cited, including the immediately contrastive use of a newly produced verb type at 2;1.

A key characteristic of Estonian, in comparison with Indo-European languages, is that it is basically agglutinative. At the same time the fact that

the inflectional alternations are generally not phonologically transparent provides a useful control for the investigation of conditions on the learning of inflections. Two of the most commonly used inflections, the third person singular past tense and the *da*-infinitive, occur in both regular and irregular forms; the set of irregular verbs is small and of high frequency. Note that of the nine different *da*-infinitive verb forms that Raivo produced, eight were also produced by Andreas; with regards to the s3past, Andreas produced seven of the eight forms that Raivo produced and an additional three that Raivo did not produce. Furthermore, the same common verbs occur in both sets of irregular forms: Of the 20 irregular verb forms produced by either child, five occurred in both s3past and *da*-infinitive in both children's first six months of verb use.

The most important difference between the two boys was the extent to which they began by learning regular vs. irregular verb forms. The focus on irregular verbs that characterized Raivo's production suggests a 'whole word strategy' or, put differently, a delay (relative to Andreas) in undertaking sub-lexical analysis. That is, Raivo appears to have begun by learning a number of high-frequency irregular verb forms (both *da*-infinitive and past tense) while failing to produce many regular forms, whereas Andreas produced most of the same irregular verb forms but additionally showed a 'spurt' in regular past tense and *da*-infinitive forms (at 1;9 and 2;0, resp.). From the analysis of Andreas' mother's speech we see that although she used far more regular than irregular past tense verb types (29 vs. 8), she used roughly the same proportion of each in tokens (.04). Thus exposure to the input should result in learning more of the frequently occurring irregular verb forms overall until the child begins to analyse out the regular past tense marker from the regular verb forms heard and used. Within the period of our study only Andreas seemed to undertake this sub-lexical analysis, which can fuel rapid verb form learning somewhat independently of input frequency.

Although the 'whole word strategy' – or failure to undertake sub-lexical analysis – could (as suggested in Vihman 1982) relate to Raivo's parallel learning of English, a language which has very few bound inflectional morphemes, at least one monolingual Estonian child has been reported to have followed the same strategy (Hendrik, also mentioned in fn. 9). Note that the figures obtained from the analysis of Andreas' mother's input speech indicate that the advantage provided by the diversity of verb types with a regular ending (but fewer tokens each) is balanced by the frequent repetition (high number of tokens) of a small number of irregular verb types – a plausible basis for an individual child to follow either learning path.

The other major differences between the two children may be related to their early attention to regular vs. irregular verbs. Specifically, we found that while Raivo began to make errors of phonological overgeneralization late and sparingly, Andreas produced larger numbers of such errors. In both cases the errors began to appear only when the child was making use of several different

verbs with both regular and irregular forms (cf. Maratsos 2000). The errors strongly suggest emergent systematicity: All show the effects of competing stem forms for the same verb type. Thus Raivo's failure to make many such errors may well be an effect of his 'whole word' strategy (well illustrated by the morphological errors). Despite the fact that he is learning increasing numbers of different inflectional forms per verb type, he does not yet seem to be making connections or generalizing across the forms of different paradigms. This may be due to the fact that although the irregular forms fall into a certain number of sub-patterns, the basic agglutinative stem + (fixed) suffix that underlies the majority of verb forms is not well modelled in the irregular verbs.

Another difference between the children involves self-reference. Andreas' morphological errors are mainly involved with his difficulty in learning which form to use for himself (see Pizzutto and Caselli 1994, for references to a large number of studies reporting the same difficulty). Second-born children exposed to English have been shown to learn the first and second person pronouns earlier than first-born children (Oshima-Takane, Goodz and Derevensky 1996); it seems likely that this is also what led Raivo to learn first and second person inflections far earlier than Andreas.

In terms of the kinds of verbal contrasts they master the two children are somewhat similar. Both Raivo and Andreas made tense, person and number contrasts and also the finite/non-finite distinction, although person was better established in Raivo's production and number in Andreas'. Andreas had established a productive affirmative/negative contrast in the past as well as the present by 2;1; Raivo had not yet discovered the productive mechanism for negative past tense expression by that age, although his present tense affirmative/negative contrast was supplemented by an invented affirmative particle. Andreas was additionally making correct use, on the whole, of the two distinct infinitives, and was productively adding new verb types in the first person plural present/imperative by the last session (recall that Andreas had also reached a higher mean length of utterance in words by 2;1 than Raivo had reached by 2;2).

In terms of the typology of verb inflection Estonian stands between maximally clear-cut morphological structures, such as that of Turkish and also Finnish, and the fusional inflections of Indo-European languages. To the extent that a language provides opaque inflectional forms with multiple functions, rote learning could be expected to dominate the learning process for a considerable period of time. Where the system is transparent, learning should be systematic and rapid. Estonian provides elements of both types; our study suggests that either learning path is possible with a mixed system of this kind.

MARILYN MAY VIHMAN
University of Wales Bangor

MAIGI VIJA
University of Tartu

5. ACKNOWLEDGMENTS

We thank those who helped us with the labour-intensive task of transcribing the recordings. For Andreas, this was a team headed by T. Hennoste, Director of the Oral Narrative project at Tartu University: Liina Lindström, Andriela Rääbis, Ave Siilak, Piret Toomet, and Ilona Tragel. For her transcriptions of Raivo's data we thank Kaja Kohler, University of Potsdam. In addition, we are grateful to Michael Tomasello, Director, Max Planck Institute for Evolutionary Anthropology, for providing the financial support for the recording and transcription of Andreas' language use from 2;0 to 2;1, and to Margaret Deuchar and Virginia Mueller Gathercole, who kindly read and commented on a draft of this paper. Finally, the Estonian Science Foundation kindly made it possible for Maigi Vija to participate in the Early Verbs conference in Berlin.

6. NOTES

[1] Raivo's mother, the first author, is a non-native speaker of Estonian, his father was a native speaker. His older sister learned Estonian as a first language and began learning English only at about age 2, but Raivo was exposed to English in half-time day-care from the age of six months as well as through his sister's English-speaking playmates.

[2] We will use Estonian orthography rather than phonemic transcription throughout: The writing system is close to phonemic; the aspects of the phonology that are not well expressed in that system are irrelevant to the issues of interest here (e.g., palatalisation, contrasts in vowel and consonant quantity which do not enter into the verbal inflection system).

[3] In over 2000 verb forms analysed, based on seven recordings of one mother's speech, seven instances were found of just one of those forms, *-mas* 'in [the process of] Xing' (*-s* marks the 'inessive', a locative case); the other forms did not occur (and neither child used any inflected infinitival forms).

[4] Only one impersonal form occurred in our child data (Andreas: *küsitakse* 'one asks', as in the adult formulaic question, *Kuidas küsitakse?* 'How does one ask?', with the expected response, *palun* 'please'; compare the corresponding English caregiver question, 'What do you say?').

[5] All of Andreas' recordings from 2;0.1 on were transcribed into CHAT format by a team trained in the transcription of oral narrative and checked by Vija, who transcribed the remaining recordings. Raivo's recordings were transcribed into CHAT format by Kaja Kohler and checked by Vihman, with the exception of the tapes from 1;9.20 and the first half of 2;1.4, both of which were transcribed close to the time of the original recording and later rechecked by Vihman.

[6] The fact that Raivo is learning Estonian in a bilingual context complicates the comparison. If his recorded English verb lexicon is included, his cumulative verb types rise to 80 by age 2;0 (adding 33 English verbs). On the other hand, there is no way to compare morphological contrast learning across the two languages. In English, he is reported in diary entries as having produced only one verb with overt contrasting inflectional forms by age 2;0: *coming, came* – but no tape-recordings were made in English situational contexts in this period.

[7] Raivo's relatively consistent provision of the copula at a time when he continues to produce many unmarked verb stems raises interesting questions in the light of Czinglar et al.'s finding (this volume) of a correlation between English, German and Croatian children's provision of the copula and their use of 'root-infinitives'. However, the best interpretation of Estonian bare stems

in relation to the 'root infinitive' is unclear, given the fact that marked infinitives are among the first forms for both children (see also Kohler 2004).

[8] As reported earlier (Vihman 1982, 1985), Raivo sometimes used English words in his Estonian utterances in this period (indicated here with small caps), although the rate of use dropped from a high of 34% in his first month of word combinations (1;8) to just 4% by 2;0.

[9] Interestingly, the invention is not unique among Estonian children. Kohler 2004 reports data from a monolingual Estonian child, Hendrik, who made use of the *jaa : ei* contrast in lieu of the s3pres inflectional marker (e.g., at 2;0.13, *tööta jaa* 'work [verb stem] yes' vs. *tööta ei* 'work no' for *töötab* vs. *ei tööta* 'works' vs. 'doesn't work').

[10] A possible source of additional confusion here is the systematic neutralisation, rare in the Estonian verbal system, between the s2 and p3 past tense forms (see Appendix).

7. APPENDIX: ESTONIAN VERBAL MORPHOLOGY

Forms used by either one of the children are indicated in **bold face**.

NON-FINITE FORMS

1st Infinitive:	(stem + *-ma*)	*katsu-ma, söö-ma* 'to [begin to, have to] try/touch, eat'
2nd Infinitive:	(stem + -[d]a)	*katsu-da, süü-a* 'to [want to, be allowed to] try/touch, eat'
Present Active Participle:	(stem + -v)	*katsu-v* 'trying/touching [adj.]', *söö-v* 'eating [adj.]'
Present Passive Participle:	(stem + -dav)	*katsu-tav* 'touchable', *söödav* 'edible'
Past Active Participle:	(stem + -nud)	*katsu-nud* 'having tried/touched', *söö-nud* 'having eaten'
Past Passive Participle:	(stem + -tud)	*katsu-tud* 'having been tried/touched', *söö-dud* 'having been eaten'

FINITE FORMS

INDICATIVE

Present	**person**	**Affix**	**Examples**	
Affirmative (pronoun) verb stem + person				
Singular	**1 (ma)**	**-n**	katsu-n	söö-n
	2 (sa)	**-d**	katsu-d	söö-d
	3 (ta)	**-b**	katsu-b	söö-b
Plural	**1 (me)**	**-me**	katsu-me	söö-me
	2 (te)	-te	katsu-te	söö-te
	3 (nad)	**-vad**	katsu-vad	söö-vad
Negative (pronoun) ***ei*** verb stem [person is unmarked]				
			ei katsu '(I, you, he/she, etc.) don't/doesn't try' ei söö '(I, you, he/she, etc.) don't/doesn't eat'	
Impersonal				
Affirmative	**-(t)akse**	verb stem + impersonal present affirmative		
			katsu-takse 'one tries' ; süü-akse 'one eats'	
Negative	- (t)a	*ei* verb stem + impersonal negative		
			ei katsu-ta 'one doesn't try, no one tries'; ei söö-da 'one doesn't eat, no one eats'	

Past					
Affirmative (pronoun)	verb stem + past + person				
Singular	**1 (ma)**	**-(si)n**		katsu-si-n	sõ-i-n
	2 (sa)	**-(si)d**		katsu-si-d	sõ-i-d
	3 (ta)	**-s, -i**		katsu-s	sõ-i
Plural	1 (me)	-(si)me		katsu-si-me	sõ-i-me
	2 (te)	-(si)te		katsu-si-te	sõ-i-te
	3 (nad)	-(si)d		katsu-si-d	sõ-i-d
Negative	*ei* + past active participle [person is unmarked]				
		-nud		ei katsu-nud ‘(I, you, he/she, etc.) didn’t try’ ei söö-nud ‘(I, you, he/she, etc.) didn’t eat’	
Impersonal	verb stem + impersonal past				
Affirmative		-ti		katsu-ti ‘people tried, attempts were made’ söö-di ‘one ate, people ate, there was eating…’	
Negative		-tud		ei katsu-tud ‘one didn’t try, no one tried’ ei söö-dud ‘one didn’t eat, no one ate’	

Present Perfect Affirmative (personal forms of present tense of verb *ole-* ‘be’ + V-nud)
Negative (*ei* + negative present of *ole-* ‘be’ + V-nud)
Past perfect Affirmative (personal forms of past tense of *ole-* ‘be’ + V-nud)
Negative (*ei* + negative past of *ole-* ‘be’ + V-nud)

IMPERATIVE

Affirmative					
Singular	1	–		–	–
	2	**- Ø**		katsu	söö
	3	-gu		katsu-gu	söö-gu
Plural	1	-gem		katsu-gem	söö-gem
	2	**-ge**		**katsu-ge**	**söö-ge**
	3	-gu		katsu-gu	söö-gu
				‘try!; let X try’ ‘eat!; let X eat’	
Negative (ära + Verb)					
Singular	1	–		–	–
	2	ära	-Ø	ära katsu	ära söö
	3	är-gu	-gu	ärgu katsu-gu	ärgu söö-gu
Plural	1	är-gem	-gem	ärgem katsu-gem	ärgem söö-gem
	2	är-ge	-ge	ärge katsu-ge	ärge söö-ge
	3	är-gu	-gu	ärgu katsu-gu	ärgu söö-gu
				‘don’t try; don’t let X try’; ‘don’t eat; don’t let X eat’	
CONDITIONAL					
Affirmative (-ks with optional person endings)				e.g., ma katsu-ks(-in) ‘I would try’ ma söö-ks(-in) ‘I would eat’	
EVIDENTIAL					
Affirmative [all persons]			-vat	katsu-vat	‘am/is/are said to try’
				söö-vat	‘am/is/are said to eat’

8. REFERENCES

Dasinger, Lisa. "Issues in the acquisition of Estonian, Finnish, and Hungarian: A crosslinguistic comparison." In Dan Isaac Slobin (ed.), *The Crosslinguistic Study of Language Acquisition*, Vol. 4. Mahwah, NJ: Lawrence Erlbaum Associates, 1997: 1-86.

Erelt, Mati, Tiina Erelt, and Kristina Ross. *Eesti keele käsiraamat [Handbook of Estonian]*. Tallinn: Eesti Teaduste Akadeemia Keele ja Kirjanduse Instituut, 1997.

Gathercole, Virginia Mueller, Eugenia Sebastián, and Pilar Soto. "The early acquisition of Spanish verbal morphology: Across-the-board or piecemeal knowledge?" *International Journal of Bilingualism* 3 (1999): 133-182.

Kohler, Kaja. Erwerb der Verbmorphologie im Estnischen. (The development of verb morphology in Estonian). PhD thesis, Potsdam University, 2004.

Maratsos, Michael. "More overregularizations after all." *Journal of Child Language* 27 (2000): 183-212.

Oshima-Takane, Yuriko, Elizabeth Goodz, and Jeffrey L. Derevensky. "Birth order effects on early language development: Do secondborn children learn from overheard speech?" *Child Development* 67 (1996): 621-634.

Pizzuto, Elena and Maria Cristina Caselli. "The acquisition of Italian verb morphology in a cross-linguistic perspective." In Yonata Levy (ed*.), Other Children, Other Languages: Issues in the Theory of Language Acquisition.* Hillsdale, N. J.: Erlbaum, 1994: 137-187.

Tragel, Ilona. "Eesti keele tuumverbidest." [On Estonian core verbs.] In Ilona Tragel (ed.), *Papers in Estonian Cognitive Linguistics.* Publications of the Department of General Linguistics 2. University of Tartu, 2001: 145-169.

Vihman, Marilyn May. "The acquisition of morphology by a bilingual child: A whole-word approach." *Applied Psycholinguistics* 3 (1982): 141-160.

Vihman, Marilyn May. "Language differentiation by the bilingual infant." *Journal of Child Language* 12 (1985): 297-324.

Vihman, Marilyn May. "The transition to grammar in a bilingual child: Positional patterns, model learning, and relational words." *International Journal of Bilingualism* 3 (1999): 267-301.

Vihman, Marilyn May and Maigi Vija. "The acquisition of Estonian locatives: A case-study using a dense developmental corpus." Paper presented at the 9th International Congress for the Study of Child Language, Madison, July, 2002.

CLAIRE MARTINOT

GRAMMATICAL ROLE OF FRENCH FIRST VERBS

Abstract. The aim of the present study is to investigate whether different functional verbs play the same role in the development of grammar and if there is a correlation between verbal grammatical marking and argument construction. The study is based on the analysis of longitudinal data from six French-speaking children aged 2;0-3;0 years. The following questions are addressed in this study: What verbal forms are prevalent in the speech of children between the ages of 2;0-3;0? Do verbs have a semantic/predicative or grammatical content? What role do verbs play in the argument construction? The analyses show that different constructions utilizing the same lexical verb are a milestone in the acquisition of syntactic structures. The data also shows that there is a different pattern of decreasing/increasing use of verb types and verb forms from the age of 2;6 on: Verb types show a decreasing tendency whereas verbal forms continue to increase. Our findings also show that French-speaking children between the ages of 2 and 3 use verbs that carry very light meaning. However, these verbs have different grammatical functions in the predication and are used with different forms of conjugation. Our analysis of the verbs most frequently used by all children in our sample enables us to claim that these verbs play an important role in the process of language acquisition by French-speaking children and can be considered language-specific. Finally, the data show that children before the age of 2;6 exhibit individual differences in lexical diversity and in their use of alternative forms of produced verbs.

1. INTRODUCTION

The notion that the construction of the utterance is syntactically determined by the verb seems to be obvious to the speaker of an Indo-European language like French, since the vast majority of French utterances are verbal constructions. The question then arises: How do French-speaking children manage to acquire proper use of this word category? On the one hand, verbs are the most variable class of words as it changes depending on person, number and tense. On the other hand, verbs are relation words used in the context of an utterance based on the fact that they may be predicative or just play an actualizing role without selecting any word in the utterance. Children construct the grammar of their language in two ways. First, they have to recognize and produce different word categories that occur in the same or different forms. Secondly, children have to combine words from different grammatical categories (nouns, verbs, adjectives, prepositions, etc.) with at least one word of another grammatical category, i.e. *preposition-noun* is possible but not *preposition-preposition*. Harris (1988) stated that the combination of words:

N. Gagarina and I. Gülzow (eds.), The Acquisition of Verbs and their Grammar: The Effect of Particular Languages, 297–318.

> contribute to the substantive information of the sentence,

i.e.,

> the partial ordering of the words is the predicational relation of operator to argument (Harris, 1988:59).

In other words, the nature of the combinations is at the same time semantic and syntactic and should not be controversial (see Ingram et al. this volume). The above-mentioned linguistic theory is totally compatible with Tomasello's (2003:93) usage-based theory of acquisition:

> learning words and learning grammatical constructions are both part of the same overall process (...) the processes of lexical and grammatical development are closely intertwined.

At a morphological level, children have to learn that some French verb forms may differ from one utterance to the other through conjugation that is through very abstract marking. This is not the case with nouns, adverbs and grammatical words as prepositions or conjunctions. However, in other languages, such as Russian, Arabic, or Hebrew, verb forms are much more numerous than in French language. Children notice as well that the meaning of nouns, adverbs (adjectives and some predicative verbs) is stable in the utterances that are directed to them. And at the same time they have to notice that a few words occur in a lot of utterances with various forms without carrying a specific meaning. These few words are very basic verbs, the most frequent in the target language and the most various on the morphological level (to be, to go, to have). We argue that children pay more attention to changes in the input than to repetitions even if they cannot understand why there is a change. Kiebzak-Mandera in this issue notices as well that children reproduce very early the alternation in pronouns they hear in the input *(I/she; you/she)*. We argue moreover that children pay more attention to changes because they correlate a change of form with a change of meaning, or information, or communicative intention (much later, from 4;0 on, they have to learn that a change of form doesn't yield systematically a change of meaning, cf. Martinot 2005:17-34). In spite of the fact that these basic verbs carry *very* little meaning per se and occur in a variety of forms, children produce them very early.

The aim of our study is to show that these few basic verbs play a grammatical role in two ways: first because children learn that the same functional category (all the verbs of the target language) occurs in various forms according to the linguistic context, secondly because the construction of the utterance and thus its meaning depends indirectly on these words.

At a structural level,

> "grammar is a set of structural regularities of language, and is produced by certain constraints on the equiprobability of words" (Harris 1988:81).

The *meaning* or the *information* is produced by "the constraints which have

> "the constraints which have the greatest effect on these probabilities, the ones that determine which words co-occur in a regular way in the utterances" (ibid.).

In the process of language acquisition a child has to discover, in the input/ intake, the instances of co-occurences, that is, which utterance is possible in the target language and which one is not or *agrammatical.* In particular, children have to notice which word selects specific ones. For example, the predicate 'to read' has to select two arguments 'I' and 'a book' in the sentence *I read a book*, but not 'my table' and 'an apple'.

The acquisition theory we have elaborated (Martinot 2000b, 2003b, 2005) attempts to explain how children produce their utterances using previously heard utterances to construct a meaning. Our theory focuses on changes which are directly connected to stages of language acquisition and reflect the level of linguistic proficiency and the cognitive/conceptual level of each individual child: a part or an aspect of the source utterance is changed while some other part or aspect remains the same. According to the reformulation theory of language acquisition we predict that at the first stage (about 0;6-1;6), children reproduce the melody of a previously heard utterance but not the same "words". At the second stage (around 1;6-2;0) children produce a few predicative and content words, identical to those they memorized and change the structure of the source utterance since they only produce one isolated word. During the third stage (2;0-4;0) children combine one content word with another word (often a word with little meaning). For example, at first a child may say *dodo* (using the noun derived from the verb *dormir* 'to sleep' which takes a predicative function); then he/she may say *fais dodo* '1S & 2S:PRES' / *fait dodo* (3S:PRES), adding the *support* verb *faire* 'to do/to make' to the predicative noun *dodo*, meaning *je/tu/il dors/t* 'I/you/he sleep(s.)'.[1] Or the child may first say *Maman* 'Mum' and then *veux Maman* 'want Mum'. In both cases the child constructs the first predicative utterance. After that he/she has to acquire all verb-governed arguments which are required in the target language: for example, $\mathbf{N_0}$ V $\mathbf{N_1}$[2] so that s/he can produce simple grammatical sentences, i.e., a sentence with one finite verb (an elementary discourse stage in children between 2;0 and 4;0 years of age).

At this third stage (children between 2;0 and 4;0 years of age), the children have the ability to retell a previously heard combination. Hence, the process of reformulation becomes more and more important because children can use both syntactical construction and addition of isolated words to construct a new utterance with new meaning. Children try to produce a lot of utterances according to the possibilities of the target language. They really construct a grammar during this third stage.

Of interest is the possible correlation between the two ways children are proceeding to construct a grammar, i.e., by increasing the use of different verb forms and by increasing the use of different verb constructions. There is data on the development of first verb forms in French language (Kilani-Schoch 2000), but not on the above-mentioned possible correlation. This paper has two aims: First, it attempts to provide evidence for such a correlation by analyzing children's speech between 2;0 and 3;0 years of age. Second, it investigates the types of verbs that mostly occur in children's utterances, i.e.:

Firstly, when the verb is predicative it governs one or more arguments, example: (2;5) *elle reste dans mon lit* 'she is staying in my bed'; and secondly, when the verb is not predicative it actualizes a noun or an adjective, as in the case in sentences with a support verb and a predicative noun, ex: (2;5) *il y avait du soleil* 'there was sun'; or adjective: (2;3) *c'est beau* 'this is beautiful'. Copulas are not predicative and can also actualize the utterance: (2;4) **l'est belle son chien* 'she is beautiful FEM his/her dog', or *(i)l est beau son chien* 'he is beautiful MASC his/her dog'.[3]

Our findings reported below show that the most frequent verbs are not full verbs but verbs with very *light* meaning. We argue that the role of the verb in the constitution of grammar doesn't necessarily result from the meaning per se of the verb, i.e., from its predicative function, but from the actualizing function of the verb in the utterance. Consequently children don't need to vary the verb in order to vary the predication or the information of utterances, e.g., they can keep the same lexical verb if a new constituent occurs in the utterance. Our findings provide support to the results presented in Ingram et al. (this volume) that the syntactical diversity of verbs is a cornerstone of language acquisition. However, our finding with French-speaking children differs from that of Ingram et al., because beginning with the ages of 2;6: Ingram's et al. findings show the increase in the syntactic diversity and verb forms, whereas our data suggest the decreasing tendency in the verb diversity.

The paper is structured in the following way: The first section provides information about French verb morphology and argument construction of sentences in the target language. The second section presents the data collected from 6 children ranging in age from 2 to 3 years old, and the method applied to obtain 2 different sets of all sentences. The first set provides utterances in the initial appearance in the speech; the second one provides sets of utterances in the alphabetical order of the finite and non-finite verb. The third section of the paper provides findings about verbal forms, which are prevalent in the speech of children. We consider both the lexical diversity (verb types) and the morphological diversity of occurring verbs. We also classify lexical forms relative to their semantic

vs. grammatical content. The last part of this section investigates the role predicative verbs play in the argument construction and further in sentence construction.

2. FRENCH VERB MORPHOLOGY AND ARGUMENT CONSTRUCTION OF SENTENCES IN THE TARGET LANGUAGE

2.1. Morphology

Many verbal forms in French language have the same pronunciation despite different pronominal values: *je/il mange* (I/he eat/s-1/3S:PRES), *tu manges* (you eat-2S: PRES), *les gens mangent* (people eat-3P: PRES), or despite different functions (infinitive *vs.* finite form) or tenses: *je veux manger* (I want to eat-INF: -*er*), *je/tu mangeais* (I/you ate, 1/2S:IPFV), *il mangeait* (he ate, 3S:IPFV), *les gens mangeaient* (people ate, 3P:IPFV), *il a mangé* (he has eaten, 3S:PFV). French language has also a variety of irregular verbs. In particular, perfect (*passé-composé*) has a very high range of variations such as *j'ai écr**it***, *j'ai l***u**, *j'ai p***eint**, *j'ai mang***é**, where the endings of the non-finite form (past participle) changes in each case.

French verbs are marked for tense: *il court* (he runs, 3S:PRES) *vs. il courait* (3S:IPFV) *vs. il courra* (3S:FUT); for person/number: *je suis*, (I am, 1S:PRES) *vs. tu es*, (2S:PRES); *je chante* (I sing, 1S:PRES), *vs. nous chantons* (1P:PRES); and for mood: *il faut* ***lire*****,** (impersonal construction with infinitive form of 'lire'= one has to read); *il faut que tu* ***lises*** (you have to read), subjunctive form; *si tu étais curieux, tu* ***lirais*** *ce papier* (if you were curious, you would read this paper), conditional form. Certain tenses in active voice and certain aspects are marked by auxiliaries (*avoir*: to have, *vs. être*: to be) before the main verb in the past participle.

Regular verbs have only one stem (je **chant**e, il a **chant**é, je **chant**erai…). These verbs have an infinitive form, ending in –**er** (chant**er**) and are the most frequent in the lexicon. However, the most frequent verbs in speech are the irregular verbs, which have numerous stems. For example, *être* 'to be' has 7 different stems, *avoir* 'to have' and *aller* 'to go' have 6 different stems, *faire* 'to do', *vouloir* 'to want', *pouvoir* 'can' have 5 different stems, *venir* 'to come', *prendre* 'to take', *savoir* 'to know', *tenir* 'to hold' have 4 different stems (see also Kilani-Schoch 2000).[4]

2.2. Argument construction

In the argument construction of verbal predicates, the subject is always obligatory and occurs mostly before the finite verb. The other arguments are a function of the verb, which commands one or two arguments with or without preposition. These occur mostly to the right of the verb.

The most frequent types of sentences in the target language (French) have three possibilities: verbs with one argument to the right of the verb, i.e. verbs with 2 arguments (N_0, N_1), verbs with two arguments to the right, i.e. verbs with total of 3 arguments (N_0, N_1, N_2) and finally, verbs with one argument, i.e. the subject (N_0). Of course it is possible to add others complements to the right or to the left of the verb, such as a temporal complement like *une fois* (once upon a time). It depends on lexical constraints and each sentence has to be examined separately.

Within the sentence, verbs can have different functions, such as predicate (1), support (2) or operator (3-4):

(1) *Pierre regarde le ciel*
'Peter looks at the sky'

But the noun can be also the predicate, in this case the verb plays a supporting role:[5]

(2) *Pierre porte son regard vers le ciel*
'Peter has (takes) a look at the sky'

Some verbs are operators on another verb in the case of the modal and aspectual verbs:

(3) *Il veut partir*
'He wants to go away'

or

(4) *Il va partir*
'He is going to go away'

3. DATA DESCRIPTION

This study has examined the use of verbs by 6 children ranging in age from 2 to 3 years old over a period of 1 year (the data for one child, Laura, are non longitudinal), see Table 1.

Table 1. The Use of Verbs (token) by Children

Child	*Total Verb count*	*2;0*	*2;3*	*2;4*	*2;5*	*2;6*	*2;8*	*2;9*	*3;0*
Laura	116				116				
Charles	108	17			31	60			
Josua	440			145	136	159			
Camille	278	25				96		157	
Thomas	238		37			108		93	
Chloé	238		134						104

As can be seen in the table, the language sample of each child was taken at different ages, and the quantity of produced verbs differs greatly across 6 children. It has to be noted, that conversations with children occurred in different situations. For example, Laura, the niece of the author, was interacting with her at night before going to bed. Charles, the son of the author, was talking to his mother in the car, at home, etc. Josua was speaking while reading picture book with his mother. The data collected from the next three children are homogeneous, i.e., Camille, Thomas and Chloé were observed while playing alone with a little house and figurines in the nursery.

The analyzed corpus contains 1418 sentences. Sentences in this sense are all the utterances with one verb in a finite form whether the sentences are well formed or incomplete. In each produced sentence, we have observed the verbal form and the argument construction to the right of the verb (see below 4. Findings).

In the following extract (Laura, 2;5), each sentence is named by the verb in the infinitive; the listing corresponds to the order of the produced speech so the continuity, repetition or rewording between the utterances are to some extent clear.

Aller; tu **vas** (rigoler)
'to go; you are going to (laugh)'
Rigoler; (tu vas) **rigoler**
'to laugh; (you are going to) laugh'
Aller; et puis Jean aussi **va** (rigoler)
'to go; and then John too is going to (laugh)'

The total number of sentences produced by each child at a given age was then reorganized in alphabetical order to obtain a second corpus, which grouped all utterances with the same lexical verbs together (extract from the corpus analyzed in alphabetical order, Laura, 2;5):

Ecouter; (je vais) **écouter**
'to listen: (I am going to) listen'
Emmener; (moi je vais) **emmener** ma poupée ma sucette
'to take: (I am going to) take my doll, my dummy'
Emmener; j'**ai emmené** ça (pour jeter à la poubelle)
'to take: I have taken this (in order to throw in the dustbin)'

4. FINDINGS

4.1 What verbal forms do French children use in their speech?

The first question we addressed in the introduction deals with verbal forms as they have been used by children in the analyzed corpus. The lexical diversity of verbs is calculated for each child at one given age. For example, (see *table 2*, Charles, 2;6) the total number of verbs in the analyzed corpus (*token*) is 60 (*table 1*); the total number of different lexical verbs (*types*) in the same corpus is 22. Therefore the lexical diversity of Charles at 2;6 is 37%.

Then we will consider the morphological diversity of all the produced verbs (1418), in other words, the number of occurrences of the same verb in different forms, as produced by the child at a given age.

In terms of lexical diversity of the produced verbs, the type/token ratio shows decreased production across age (*table 2*) The youngest children have the highest ratio of lexical diversity in the use of verbs. In spite of the fact that the lexical diversity of the produced verbs shows decreasing tendency in 3 children (Charles, Camille, Thomas) across ages, the same tendency was not observed in sentences produced by Josua and Chloé, which leads us to postulate that the lexical diversity is probably more child specific than age specific. Nevertheless the child specificity is much more obvious before the age of 2;6.

Table 2. The Diversity of Verbs Across Ages (%)

Child	*2;0*	*2;3*	*2;4*	*2;5*	*2;6*	*2;8*	*2;9*	*3;0*
Laura				32				
Charles	94			58	37			
Josua			23	27	26			
Camille	40				27		25	
Thomas		40			22		29	
Chloé		24						24

On the other hand, if one considers the production of alternative forms for each verb, i.e., same verb produced in 2, 3 or 4 different forms, there is a clear increase in production as a function of age, see table 3:

Table 3. Rate (%) of Verbs Occurring in Several Forms Across Ages

Laura	*Charles*			*Josua*			*Camille*			*Thomas*	
2;5	2;0	2;5	2;6	2;4	2;5	2;6	2;0	2;6	2;9	2;3	2;6
30	0	5	**39**	24	11	**29**	10	24	**41**	13	**25**

Once again, in spite of the very clear increasing tendency of the production of alternative forms for each verb in the 4 children above, the comparison of rates at the same age shows a big difference suggesting that the morphological development is child specific.

However, there is a correlation between the decreasing lexical diversity of the verbs relative to age (*table 2*) and the increasing use of different verbal forms (*table 3*). The increasing ability of a child to use different forms of verbs corresponds to one of the grammatical marking processes. One can consider that grammatical marking (which carries no meaning for a child at this age) puts verbs in a less meaningful category (unlike nouns, for ex.) and thus in a more grammatical category with abstract marking. Then children can use verbs or some verbs, the most frequent in fact, as pure grammatical units. In playing a more grammatical role verbs can be used in more than one construction (see below ex. 27-30). (The same situation is observed in a lower measure in the target language). Thus, children have to use the same verb that has different lexical meaning according to the situation (*être* has a different meaning or value in : *il est à la maison* 'he is at home' and *il est heureux* 'he is happy'). In these cases, different meanings correspond to different sentence constructions. In this view morphological marking would be a condition for the verb to determine the construction pattern of an utterance. If the verb can be used in 2 or 3 different constructions yielding 2 or 3 different utterance meanings, then the need of new verbs is no more so drastically urgent. The syntactical development in children (with several constructions for the same verb) could explain the decreasing tendency to verbal diversity from 2;6 on.

The increased ability of a child to use different forms of verbs means that the child is able to use the specific grammatical properties of a word class, verbs in our case, in the construction of a predication.

The most important observable factor in the correlation between the decreased lexical diversity and increased use of different verbal forms, is the way children seem to use verbs: first as units describing sense (stage a, at the beginning of the above-mentioned third stage), and later, at around the age of 2;6, as sense *and* grammatical units (stage b).

Actually, the meaning of the same verbal form in several different sentences is determined by syntactic structure. Therefore, it is possible to postulate that grammar allows the child to construct different sentences with different meanings, without increasing the lexical variety. The lexical diversity of the verbs can consequently decrease without compromising the construction of meaning. At the same time, the variation of the morphological forms determines, among others, the way an utterance is actualized in a language like French.

The list of the verbs occurring in different forms used by at least one child (2;0-3;0) is represented in figure 1.

Figure 1. Classification of Verbs Occurring in Different Forms

a. **être – aller – avoir** 'to be, to go, to have' – verbs with little meaning, used as operators (aller) or as support verbs (être, avoir).
b. **faire – mettre – prendre – donner** 'to make/to do, to put, to take, to give' – verbs with little meaning, often employed as support verbs.
c. **tomber** – **rentrer** – **sortir** – **avancer** – **courir** – **rouler** – **se promener** – **monter** – **emmener** 'to fall; to go in; to go out; to go on; to run; to roll; to walk; to go up; to take with' – motion verbs.
d. **manger** – **dormir** – **se coucher – (se) réveiller** – **jouer** – **fermer** – **ouvrir** – **toucher** 'to eat; to sleep; to go to bed; to wake up; to play; to close; to open; to touch' – verbs denominating very common children's activities.
e. **voir – regarder** ' to see; to look at' – verbs of perception.
f. **dire** 'to say / to tell' – **parler** 'to speak' – **raconter** 'to tell' – **téléphoner** 'to phone' – **s'appeler** 'to be the name of' – verbs referring to speech activity.

The first three verbs, as listed in (a), are the most frequently used in speech of all children from the age 2;6 on. The next four verbs, as listed in (b), although frequent, appear less systematically than the first three. The verbs, as listed in (c-d-e-f), are much less frequent than the verbs as listed in (a-b). Those verbs have specific meaning. The verbs as listed in (a-b) can be manipulated for grammatical purposes.

The occurring forms are seen in regular verbs ending with '-e' (1S & 3S:PRES: *mang**e***), '-es' (2S:PRES: *mang**es***), '-ent' (3P:PRES: *mang**ent***), '-é'

or '-er' (respectively past participle: *mang**é*** 'eaten' in the compound form, perfect: *a mang**é*** 's/he has eaten' and infinitive: *mang**er*** 'to eat').

The occurring forms of irregular verbs used by children appear in the same tense, but they are easier recognizable and can be divided into infinitive form (*dormir*: 'to sleep'), simple finite form in the present tense 1S, 2S, 3S (*dors/t*: 'I, you, s/he sleeps'), and compound inflected form (*as/a dormi*: 'you have / s/he has slept'). The ratio of irregular verbs[6] to regular verbs used by all children in the sample across all ages is quite stable at 49%. On the average, 90% of the verbs in a simple finite form are used in the present tense or more rarely in imperfect.

We have seen that the lexical diversity of the produced verbs decreases at around the age 2;6. At the same time, the rate of verbs' occurrences in several different forms increases. We assume that at around the age of 2;6 a few verbs produced by each child have either a predominantly semantic role or a predominantly grammatical role in the predication. For the most part, the verbs occurring in different forms are those that have little meaning, i.e., supporting verbs, and few operators, such as ***aller*** in *je **vais** me lever* 'I **am going to** stand up'. Thus, the morphological markings are seen mostly in verbs, which have a predominantly grammatical role in the predication (see *fig.1* in verbs as listed in a-b).

4.2. Semantic/predicative or grammatical content?

The second question we addressed in the introduction investigates the role of the verb of each sentence: does the verb have a full meaning (predicative content) or it has little meaning and is used for grammatical content? In the first case, the verb selects the arguments (predicative role), in the second case the verb introduces any verb: for ex. *vouloir* 'to want' is an operator in: *je veux (manger + dormir + partir…)*: 'I want to (eat + sleep + go away…)', or is selected by the predicative noun or adjective in: *il met la table* 'he sets the table', *met* is the support verb appropriated to the noun *table*, *vs. *il fait la table*; in: *il fait beau* 'the weather is beautiful', *fait* is the support verb in the impersonal construction appropriated to the adjective *beau, vs. *il est beau* 'he is beautiful'.[7] Consequently, the meaning of the verb only plays a role in the argument construction when the verb itself has predicative content. The sentence construction is thus organized according to the meaning of the verb. When the verb is a support verb, it doesn't play any role on the argument constraints, but, as a verb, it plays a role in the predicative construction of the sentence by actualizing the predicative noun or adjective. That is why we argue that support verbs may be considered grammatical words, similar to the role of prepositions.

If we compare verb frequency from 2;6 on, it is very clear that the most frequent verbs are the 'grammatical' verbs, in particular *être* 'to be' and *aller* 'to go', (see *table 4*).

The verb *être* normally occurs in representative way as a support verb in sentences like:

(5) *c'est Marie*
'this is Mary'

or as a copula in sentences like :

(6) *ils sont gentils*
'they are nice'

aller occurs as an aspectual operator :

(7) *elle va le mettre là*
'she is going to put it there'

less often as a motion verb:

(8) *on va là*
'we go /are going there'

avoir 'to have' occurs as a support verb in sentences like:

(9) *il y a une chaise*
'there is a chair'
(10) *j'ai peur, j'ai froid*
'I am afraid, I am cold'

faire 'to do, to make' occurs as a support verb in sentences like:

(11) (N_0) *fait des sourires*
'N_0 does smiles = smiles'
(12) *il fait nuit*
'it is dark'

or less often as an operator:

(13) *tu les fais manger*
'you make them eat'

mettre and *prendre* 'to put, to take' are divided up in support verbs like:

(14) *il met la table*
'he sets the table'
(15) *il prend un bain*
'he has a bath'

and predicative verbs like:

(16) *il met la main sur la table*
'he puts his hand on the table'
(17) *il prend son chapeau*
'he takes his hut'

Table 4. Frequency (%) of Verbs 'être' and 'aller' (%) at Different Ages

	Laura	*Charles*		*Josua*			*Camille*			*Thomas*			*Chloé*	
	2;5	2;0	2;5	2;6	2;4	2;5	2;6	2;0	2;6	2;9	2;3	2;6	2;9	2;3	3;0
Être	20	0	19	13	34	41	24	0	21	29	0	29	28	18	32
Aller	12	0	0	15	4	0	9	0	18	14	11	0	4	19	19
Tot	32	0	19	28	38	41	33	0	39	43	11	29	32	37	51

Our data shows that on the average, one third of all utterances produced by children before 3 years of age have *être* or *aller* as finite form (*table 4*). However, one can notice that at the same age (2;5) Josua produces twice as many occurrences of *être compared* to Laura and Charles, whereas two boys (Josua and Charles) didn't produce the verb *aller* at all compared to Laura. At 2;6, Thomas doesn't show any use of *aller,* whereas Camille produces twice as many verbs *aller* than Josua.

Among the relatively frequent verbs, as listed in (b), only *mettre* 'to put' and *prendre* 'to take' have a predicative function in some sentences.

4.3. What role do predicative verbs play in the argument construction and sentence construction?

Now we have to consider how the predicative verbs are used by children between the ages 2;0 and 3;0 to select their arguments. Based on our data, it is possible to show that predicative verbs play an important role in the selection of nouns to the right of verb. In speech, children add another complement (a reformulation procedure):

(18) Laura (2;5)
**Je parte* (= pars) *la nuit*
'I go away in the night'
Je parte* (= pars) ***en vacances* *la nuit*
'I go on holiday in the night'
Elle reste là la poupée
'it is staying there, the doll'
Elle reste ***dans mon lit*** (prep N_{loc})
'it is staying in my bed'

They also try to change the order of the constituents (another reformulation procedure):

(19) Charles (2;6)
En plus *moi je tousse comme Bapou*
'in addition to that I cough as Bapou'
Je tousse pas ***en plus***
'I do not cough in addition to that'

They add a negation (another reformulation procedure):

(20) Josua (2;4)
Celui-là il pleure
'this one, he is crying'
Il ***pleure pas*** *celui-là*
'he does not cry, this one'
**L'est belle son chien*
'(she) is pretty, her dog'
Elle est* ***pas belle* *celui-là*
'she is not pretty, this one'
Ouais l'est ***pas belle***
'yes, (she) is not pretty'

However, the most frequent sentence constructions at that age range appear to be the ones that consist of repeating the verb and changing the noun (another reformulation procedure) as if children were trying all possibilities to the right of the verb. It is important to note again that this very productive procedure applies most often to verbs which are not predicative:

(21) Laura (2;5)
*Jean il était **en vacances***
'John he was on holiday'
Il était **dans la salopette*
'he was in the dungarees'
*Il était **dans mon panier***
'he was in my basket'
*Il était **dans son lit***
'he was in his bed'

Josua (2;8)
*Elle est pas **gentille** Cruella*
'she is not kind, Cruella'
*Maintenant elle est plus **méchante***
'now she is no more nasty'

Josua (2;8)
*C'est **le bébé**?*
'this is the baby?'
*C'est **la police**?*
'this is the policeman?'
*C'est **le méchant***
'this is the bad one'
*C'est **le premier***
'this is the first one'
*C'est **Pongo*** 'this is Pongo'

Laura (2;5)
*Il y avait du **tonnerre***
'there was thunder'
*Il y avait du **soleil***
'there was sun'

We argue that children need much more time and have to reach a higher level in their language development to be able to produce sentences, which are organized by a verbal predicate than sentences in which the verb only actualizes a noun.

The last question to be addressed is the possibility of finding any correlation between the alternation of verbal forms and the syntactical construction.

Let us consider the utterances of Thomas and Chloé (appendix) who have very different productions at 2;3 years of age respectively. Thomas shows a

contrast between the finite simple form and the infinitive form but no correlation between the verbal form and the argument construction. For example, the verb *dormir* 'to sleep' occurs with one form, *dort*, and two different constructions, the intransitive (22) and the locative one (23):

(22) **dors/t (petit bébé + nounours)*
'sleep(s) little baby + teddy bear'

(23) **dort dans chambre*
'sleeps in bedroom'
**dort là*
'sleeps there'

The subject does not occur systematically (not a possibility in French) without any alternation of the verbal forms. However, the choice of arguments only concerns the subject:

(24) **dort petit bébé*
'sleeps little baby'
**dort nounours*
'sleeps teddy bear'
**il se promène monsieur*
'he takes a walk Sir'
**promène nounours*
'takes a walk teddy bear'

In all Thomas's verbs, with and without formal alternation, there is only one argument in position N_1 (prepositional):

(25) **va aller à la voiture (nounours + celui-là)*
'is going to go to the car (teddy bear + this one)'

and only one locative prepositional complement:

(26) **dort dans chambre*
'sleeps in bedroom'

The only real contrast in the construction appears with *aller* as a motion verb (infinitive form):

(27) **va **aller** à la voiture*
'is going to go to the car'

and as an operator:

(28) ***va** pas promener*
'is not going to walk'

Now let us look at Chloé's productions: The verb *aller* 'to go' occurs in 4 different forms which are the infinitive (*aller*), the simple present form in the 1S (*je vais*), 3S (*il va*) and 3P (*ils vont*). However, she uses the same constructions as Thomas with *aller,* i.e., as an operator in the following sentence:

(29) *il **va** aller dans le parking*
'he is going to go into the car park'

She also uses *aller* as a motion verb with 3 different prepositions and a large choice of locative arguments such as:

(30) *il va **aller** jouer **dans la maison de moi***
'he is going to go and play into the house of mine'
*il va **aller chez Bilette***
'he is going to go to Bilette ('s home)'
*ils **vont** pas **au gymnase***
'they do not go to the gymnasium'

It seems that at the age of 2;3 the correlation between the alternation of verbal forms and the syntactical construction is child specific.

However, these examples indicate that children attempt to use all the possible nouns/verb combinations they know as if they were trying to experiment with the co-occurrence constraint. In the case of the verb *aller* 'to go' it is possible to see a relationship between the formal alternation and the diversity of the locatives, at least in Chloé's production.

The verb with the most numerous contrasts in the construction but not necessarily in the formal alternation is the verb *être* 'to be' which occurs only with *est* (3S:PRES) or *sont* (3P:PRES). As a support verb, it introduces quantity: *c'est beaucoup* 'this is a lot'; it introduces a pronoun: *c'est ça* 'it is this'*, c'est moi* 'that's me'*, c'est quoi* 'what is this?'; a noun: *c'est pas une chaise* 'this is not a chair'; a prepositional complement: *c'est comme Helda* 'this is like Helda'*, c'est pour celle-là* 'this is for this one'*;* and an adjective: *c'est plus facile* 'this is easier'. As a copula, it introduces an adjective: *les bonhommes ils sont trop petits* 'the little men they are too little' and an adverb: *ils sont debout* 'they are up'.

The last example of children's verb constructions concerns the use of the verb *manger* 'to eat'. This verb occurs in 3 different forms, such as infinitive, simple present and compound forms. It is also used in 2 different constructions: 1) the intransitive (*faut manger*- impers. form for *you have/he has to eat*) and 2) the transitive. The transitive form can occur with 3 different complements, such as: *ils vont manger* ***le dîner*** 'they are going to eat the dinner'*, celui-là il mange* ***des pâtes*** 'this one he is eating pasta'*, il a* ***tout*** *mangé* 'he has all eaten'. In the noun category, *le dîner* 'dinner' refers to the things we eat during the dinner, *des pâtes* 'pasta' refers to a particular food.

5. CONCLUSION

The data presented in this study shows that there are specific functions that verbs serve in a speech production of children between 2;0 and 3;0 years of age. In our sample the verbs were used in a variety of ways: as predicate or support; operator or predicate; copula or predicate. The increased verb occurrences appear to be good predictors for language acquisition. Moreover, different syntactic properties can be expressed with the same functional verbs used in different occurring constructions. These will include the presence or absence of the argument in position N_1 (*je mange* vs. *je mange une pomme* – 'I eat' *vs.* 'I eat an apple'), or the presence or absence of another complement (*je vais à l'école* vs. *chaque jour je vais à l'école* – 'I go to school' *vs.* 'every day I go to school'). The syntactic properties of the abovementioned verbs and their flexibility to be used with different functions in the predication are verb-specific as well as language specific.

The lexical diversity of all nominal arguments appear to be a good predictor for language acquisition as well, because the lexical choice determines if the construction is grammatical or not, or makes sense or not (ex. 21,30).

In conclusion, there is evidence for a correlation between an increase of verb forms and increase of different argument construction only for the most frequent verbs which are very few in numbers in child's speech before three years of age and have very little meaning in French language. However the analysis of our data supports the fact that all verbs show an increase in the use of different forms for the same lexical verb across ages and indicates that the variety of the argument construction or the functional variety of some verbs is a good predictor for the grammatical development. Children have to learn which morphological features, that are which possible forms in a given context, can be used in verbs. Then they have to learn which arguments co-occur with each verb in order to produce a grammatical sentence and they have to learn that, relative to linguistic context, some verbs can have a predicative function - a semantic one, or a grammatical function (as operator,

copula, auxiliary or support). Both morphological development and the different use of a verb construct a decisive part of the grammar.

Nevertheless, this evidence can not be extended to all children. Our results show that before the age of 2;6 the lexical diversity of produced verbs and the rate of production of two grammatical verbs that are most frequently used on the average are child specific. In all ranges of age between 2;0 and 3;0, the production of alternative forms for each discussed verb seems to follow the same child-specific pattern.

CLAIRE MARTINOT
Université René Descartes, Paris V
LEAPLE, UMR 8606

6. NOTES

[1] The *support* verb *faire* allows the predicative noun *dodo* to take a function of a verb. The main feature of *support* verbs *(*some of the so-called *light* verbs) is that they are reducible in certain restructurations.

[2] The V represents a finite verb, N_0 the subject and N_1 the argument of range 1.

[3] Ibrahim (2000) has proposed 6 different levels of verb classification based on the degree of appropriation the verb holds in relation to the next word.

[4] The conjugation of French verbs are commonly distinguished according to their infinitive form and can be classified into 3 groups: 1) conjugation group consisting of infinitive form ending with *–er*; 2) conjugation group that has infinitive form ending with *–ir*, like *finir*, and has regular forms and only one stem; and 3) conjugation group that consists of all irregular verbs which have several stems.

[5] This category of verbs was discovered by Harris (1964) to describe the nominalization of: *he walked* into *he took a walk*. In this transformation, the verb *to take* has only a grammatical function and does not express the predicative meaning of the sentence. In German, this very important category of verbs is named *Funktionsverben* (von Polenz 1963). In French, they are named *verbes supports* (Daladier 1978) and are described by many researchers (Giry-Schneider 1987; Gross 1976, 1981; Ibrahim 1994, 1996, 2000)

[6] As mentioned above, those verbs that are classified as irregular have more than one stem in the conjugation.

[7] The sentences: "*il fait la table"* and "*il est beau"* are agrammatical if used in the same context as "*il met la table"* and "*il fait beau",* respectively.

7. APPENDIX

Construction types for verbs which occur more than one time in 2 children of 2;3 years of age

Thomas
1a. va aller à la voiture (Nounours + celui-là)
1b. va pas promener
2a. dors/t (petit bébé + nounours)
2b. dort dans chambre

2c. dort là
3. fait tomber (bonhomme + le monsieur)
4a. il mange
4b. là il mange / là mange
4c. mange là nounours
4d. tout mangé
5a. non va pas promener celui-là
5b. il se promène monsieur
5c. promène nounours
6a. fait tomber (bonhomme + le monsieur)
6b. tombE (bonhomme + le monsieur)

Chloé
1a. il va aller (dans le parking + là)
1b. il va aller jouer dans la maison de moi
1c. il va aller chez Bilette
1d. ils vont au (jardin + marché)
1e. ils vont dans le parc
1f. ils vont pas au gymnase
2a. il y a pas (de piles + une chaise)
2b. il y en a beaucoup – encore il y en a plus
2c. ils ont cuillère pour le lapin aussi
2d. ils ont pas de mains
3a. faut donner une cuillère pour manger
3b. tu donnes ça
4a. c'est beaucoup – c'est tout
4b. c'est ça – c'est moi – c'est quoi
4c. c'est pas (une chaise + une table + un camion + un petit tabouret)
4d. c'est comme Helda – c'est pour celle-là
4e. c'est plus facile
4f. *celle-là ils sont beaux –celui-là il est (noir + gentil + pas gentil + très dangereux) – les bonhommes ils sont trop petits
4g. ils sont debout
4h. ils sont arrivés
5a. vont faire (des galipettes + les courses) – font pas dodo
6. faut (pas + E) V_{inf}
7. ils ont (pas + E) fini – fini de ranger
8a. faut manger – il mange
8a'. ils vont manger debout
8b. ils vont manger le dîner
8b'. celui-là il mange des pâtes
8b''. il a tout mangé
9a. faut mettre les joujoux dedans
9b. je vais te mettre debout
9c.* je mets – *on met les enfants
10a. faut (pas + E) les ranger (les fourchettes + les cuillères)
10b. fini de ranger
11a. ils roulent
11a'. pour rouler avec des chaises
12. ça va tomber – il tombe pas le camion – lui il est tombé
13. va pas toucher mes bonhommes – ne touche pas mes bonhommes – tu as pas touché mes bonhommes.

8. REFERENCES

Bloom, Lois. *Language Development: From two to three*. Cambridge: Cambridge University Press, 1991.

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

Catell, Ray. *Composite Predicates in English, Syntax and Semantics*. New York: Academic Press, 1984.

Daladier, Anne. "Quelques problèmes d'analyse d'un type de nominalisation et de certains groupes nominaux français." Thèse de 3ème cycle. Université de Paris 7, 1978.

Giry-Schneider, Jacqueline. *Les prédicats nominaux en français : les phrases simples à verbes supports*. Genève: Droz, 1987.

Gross, Maurice. "Sur quelques groupes nominaux complexes." In Jean-Claude Chevalier & Maurice Gross (eds.), *Méthodes en grammaire française*. Paris: Klincksieck, 1976: 97-119.

Gross, Maurice. "Les bases empiriques de la notion de prédicat sémantique." *Langages* 63 (1981): 7-52.

Gross, Maurice. "La structure d'arguments des phrases élémentaires." *Lingua Franca* 1 (1994): 7-17.

Harris, Zellig S. *Elementary Transformations*. T.D.A.P. (Transformational and Discourse Analysis Papers). University of Pennsylvania, Philadelphia, 1964.

Harris, Zellig S. *Language and Information*. New York: Columbia University Press, 1988.

Ibrahim, Amr H. "Théorie générale: de la nature du support de neutralité articulant l'arbitraire et le motivé dans les langues." In Amr Ibrahim (ed.), *Supports, Opérateurs, Durées*. Annales Littéraires de l'université de Besançon 516, Paris: Les Belles Lettres, 1994: 15-61.

Ibrahim, Amr H. "La forme d'une théorie du langage axée sur les termes supports." *Langages* 121 (1996): 99-119.

Ibrahim, Amr H. "Une classification des verbes en six classes asymétriques hiérarchisées." In Françoise Cordier, Jacques François & Bernard Victorri (eds.), *Syntaxe et Sémantique* 2. Caen: Presses Universitaires de Caen, 2000: 81-97.

Ingram, David. *First Language Acquisition*. Cambridge University Press, 1989.

Kilani-Schoch, Mariane. "Early verb development in two French-speaking children." *ZAS Papers in Linguistics* 18 (2000): 79-97.

Martinot, Claire. "Prédicats et supports chez un enfant de 3 ans." *Langages* 121 (1996): 73-91.

Martinot, Claire. "Les verbes supports dans l'acquisition de la syntaxe." In Claire Martinot (ed.), *Actes du Colloque International de Besançon sur l'acquisition de la syntaxe en langue maternelle et en langue étrangère.* Annales littéraires de l'Université de Franche-Comté, n° 631, Paris: Les Belles Lettres, 1997: 87-106.

Martinot, Claire. "Développement de la construction argumentale de trois verbes essentiels : mettre, prendre, donner." *Langue française* 118 (1998): 61-84.

Martinot, Claire. "Deux phénomènes de grammaticalisation dans l'acquisition du français langue maternelle." In Annick Englebert, Michel Pierrard, Laurence Rosier & Dan Van Raemdonck (eds.), *De la grammaire des formes à la grammaire du sens.* Tübingen: Max Niemeyer, 2000a: 345-352.

Martinot, Claire. "Etude comparative des processus de reformulations chez des enfants de 5 à 11 ans." *Langages* 140 (2000b): 92-124.

Martinot, Claire. "Acquisition des verbes et reformulation." In Karine Duvignau, Olivier Gasquet, Bruno Gaume (eds.), *Revue d'intelligence artificielle*, vol.17 (2003a) : 787-798.

Martinot, Claire. "Les acquisitions tardives en français langue maternelle." In Claire Martinot, Amr H.Ibrahim (eds.), *La Reformulation: un principe universel d'acquisition*. Paris :Kimé, 2003b : 39-69.

Martinot, Claire. *Comment parlent les enfants de 6 ans. Pour une linguistique de l'acquisition*, Presses Universitaires de Franche-Comté, 2005.

Olguin, Rachel and Michael Tomasello. "Twenty-Five-Month-Old Children Do Not Have A Grammatical Category Of Verb." *Cognitive Development* 8 (1993): 245-272.

Tomasello, Michael. *First Verbs: A Case Study of Early Grammatical Development.* Cambridge: Cambridge University Press, 1992.

Polenz, Peter von. *Funktionsverben im heutigen Deutsch. Sprache der rationalisierten Welt.* Wirkendes Wort, Beiheft 5. Düsseldorf: Schwann, 1963.

DOROTA KIEBZAK-MANDERA

SPEAKER AND HEARER REFERENCE IN RUSSIAN SPEAKING CHILDREN

The complexity of personal pronoun distinction resides more in the head of the psycholinguist than in the head of the child.

Chiat 1986:354

Abstract. The purpose of this research[1] is to determine how Russian speaking children code the distinction between the speaker and the hearer. I refer to the corpora of two children: a girl, Varja and a boy, Kirill. The children's data are compared with the data of their mothers' speech. It is well known that the speaker and the hearer can be marked in two ways. One can use either standard forms of 1:SG and 2:SG (e.g. *Chego ty vyp'jesh'?* 'What will you drink?') or instead of them – forms of 3:SG (e.g. the mother says to Varja: *Varja chego vyp'jet?* 'What will Varja drink?'). These two trends were noted in the analysed material. Kirill's mother hardly ever used forms of 3:SG when marking the speaker and the speech addressee. Moreover, these forms were used rather on special 'solemn' occasions. A low percentage of these forms is observed in all the material. Similar results were noted in the boy's corpus. The analyses of Varja's and her mother's data lead to completely different results. They both treated the forms of 3:SG as one of the normal ways of marking the speaker and the speech addressee. The percentage of those forms – both in the girl's and her mother's data – was high at the early stages, but it immediately dropped when Varja was about 2;0. As for the personal forms, it was only Varja who had problems with the proper use of them. She tended to use forms of 1:SG in order to mark the speech addressee and forms of 2:SG in order to mark the speaker. This phenomenon was limited in time and probably connected with the specific language-learning strategy adopted by the girl.

1. INTRODUCTORY REMARKS

In Russian, the category of person is marked in the forms of all three moods: indicative, imperative and conditional. The analysis is based, however, only on non-past tense forms of the indicative. Among imperative forms, the forms of 2:SG are predominant – they constitute over 90% of all usages. Conditional mood appears in children relatively late and is very infrequent: it constitutes only about 1% of all verbal forms.

In the present and future perfective tenses in Russian, personal endings are attached directly to a verb's stem. Future imperfective forms are periphrastic: here, an auxiliary verb is a carrier of the category of person. So, personal forms of non-past tenses, marked by overt, specific endings are so

N. Gagarina and I. Gülzow (eds.), The Acquisition of Verbs and their Grammar: The Effect of Particular Languages, 319–344.

clear and unambiguous that a presence of the subject in the clause is not always necessary and sometimes even redundant (cf. *Grammatika* 1980).

In the past tense, the opposition between personal forms is neutralised,[2] so the forms are not formally varied. Lack of a subject or the impossibility to deduct it from the context make it sometimes totally impossible to determine whether a certain past form used in an utterance refers to the speaker, the addressee or a third person. This dilemma often appears during the analysis of material which consists of transcribed utterances with only few situational comments – as was the case in the data analysed in the present study.

The ways of marking the person in the indicative in Russian are presented below:

Table 1. Person in the indicative mood in Russian

	imperfective form		*perfective form*	
PRESENT	imperfective present tense stem + personal endings			
	SG:1	**pish-u**		
	SG:2	**pish-esh'**		
	SG:3	**pish-et**		
	PL:1	**pish-em**		
	PL:2	**pish-ete**		
	PL:3	**pish-ut**		
FUTURE	auxiliary *byt'* ('to be') in personal form + imperfective infinitive		perfective present tense stem + personal endings	
	SG:1	**bud-u pisat'**	SG:1	**napish-u**
	SG:2	**bud-esh' pisat'**	SG:2	**napish-esh'**
	SG:3	**bud-ut pisat'**	SG:3	**napish-et**
	PL:1	**bud-em pisat'**	PL:1	**napish-em**
	PL:2	**bud-ete pisat'**	PL:2	**napish-ete**
	PL:3	**bud-ut pisat'**	PL:3	**napish-ut**
PAST	imperfective past tense (infinitive) stem + suffix -l- + gender endings ○ neutralisation of person in singular and plural ○ neutralisation of gender in plural		perfective past tense (infinitive) stem+suffix -l- + gender endings	
	SG:M	**pisa-l**	SG:M	**napisa-l**
	SG:F	**pisa-l-a**	SG:F	**napisa-l-a**
	SG:N	**pisa-l-o**	SG:N	**napisa-l-o**
	PL	**pisa-l-i**	PL	**napisa-l-i**

1.1. Data

The paper is based on the data of two Russian-speaking children: a girl, named Varja and a boy, named Kirill. Varja was observed from 1;2 to 3;0. Her data consist of a speech diary of about 300 type-written pages and seven tape-recordings. The latter were made at the ages of 1;6, 1;7, 1;8, 1;10, 2;0, 2;4 and 2;10. The material was gathered and made available by Ekaterina Protassova. After having done provisional calculations of Varja's material, I have put together the data of both sources coming from the same age: a bigger volume of the material from some months did not impact its structure (cf. Kiebzak-Mandera 1996). Thus, in Varja, I identified a total of 9763 verbal forms, out of which 4906 were used for the present study.

Kirill's data come from tape-recordings only and they encompass a period of 1;8-3;0. They were collected and transcribed by Elena Bieva.[3] In the boy, I identified a total of 6340 verbal forms, out of which 2963 were used for the analysis. Children's data were compared with the input language, i.e. with their mothers' corpora. In Varja's mother, 3810 verbal forms were registered, out of which 1818 were analysed here. In the boy's mother's data, 14677 and 7246 forms were registered and analysed respectively. The quantitative data of the analysed material are presented in Table 2:

Table 2. Number of verbal forms in the analysed corpora

	all registered verbal forms	*analysed verbal forms*
VARJA'S DATA	9763	4906
KIRILL'S DATA	6340	2963
VARJA'S MOTHER'S DATA	3810	1818
KIRILL'S MOTHER'S DATA	14677	7246

All the data have been entered into a computer and elaborated according to the *CHILDES* system (MacWhinney 1991). Varja's and Kirill's language development in the early stages proceeds in quite a different way. Varja was, with no doubt, a so-called *phrasal child* (see Pine and Lieven 1993, cf. also Bates, Dale and Thal 1995: *an early talker*). Such children first remember whole long expressions, use them in adequate situations, analogical to those in which they had been heard, and only after some time they begin to analyse utterances by decomposing them into components. Even if the speech of such children seems to be well-developed from early on, many of the structures they use are not productive[4] yet.

However, not all of the early verb forms used by the girl should be regarded as the components of the memorised formulae or as amorphic,

"fossilised" forms (typical for the early stage of child language development, cf. e.g. Gvozdev 1949). Varja started using various forms built on the base of the same stem very early – and in a lot of cases, their usage was clearly creative.[5] The turning point in the development of verbal morphology in Varja was after 1;5; then formulae began to disappear in her speech too.

As for the boy, only a few verbal forms could be observed in the early recordings between the ages of 1;8-1;9. The turning point in his language development was at 1;11.[6]

The results of the analyses of the earliest stages of the children's language development, when the verbal morphology was being formed, are presented in monthly intervals. Later data, encompassing the period after the frequency of particular verbal categories was stabilised, are presented in quarterly intervals. Quarters are marked with Roman numerals, the quarter's number stands for the actual quarter of the child's life.

1.2. Acquisition of person category by children – state of the art

The majority of publications on the acquisition of the person category by children focuses on English. In English, the person is very rarely marked with an ending attached to a verbal stem. The researchers, therefore, have concentrated on the process of acquisition of personal pronouns as the main determinant of the category of person. For example, in Chiat's (1986) opinion, children begin rather with pronouns referring to the speech addresser, 1:SG forms, thereafter they start using forms of inanimate 3:SG pronouns and later on – pronouns referring to the speech addressee, 2:SG. According to Chiat, children have no problems with the personal pronouns: the roles in a speech act are so obvious that probably children get accustomed to them even before they start speaking.

Slavic data were analysed, among others, by Smoczyńska (1992): she investigated the corpora of five Polish children. Smoczyńska states that the reference to the speaker appears in all children approximately at the age of 1;6. From the very beginning children mark the speaker by forms of 3:SG, but forms of 1:SG also appear in their data, although definitely more rarely. 1:SG forms start to appear more frequently only at the beginning of the third year of children's life: about the age of 2;3 this way of marking the speaker is predominant. Smoczyńska notices that 2:SG forms used to mark the hearer appeared in the children's data either simultaneously with 1:SG forms or a bit later.

2. REFERRING TO SPEAKER AND HEARER IN THE INPUT LANGUAGE

Adults, referring to themselves and to their children, use one of the two strategies. They mark the speaker and the hearer in the standard way by forms of 1:SG and 2:SG respectively:

(1) K. 1;9[7] M: *Uberi svoju mashinu, ty zadavish' ljalju, ja ejo sjuda perestavlju.*
'Take your car away, you will strangle the doll, I will put it there.'

Alternatively, they resign from marking personal distinctions in favour of one, universal, 'neutralised' form of 3:SG, while the reference is expressed by a noun subject: *mummy* when referring to the speaker, and child's name when referring to the hearer:[8]

(2) K. 1;10 M: *A chto mama risuet?*
'And what is mummy drawing?'
K: *Afisju (= avtobus).*
afisja:SG:ACC
'A bus.'
after a while:
M: *Vot mama kakoj xudozhnik.*
'What a painter mummy is.'
M: *Tak, nu a Kirjusha chto narisuet?*
'Yes, and what will Kirjusha draw?'

They are hardly ever consistent, however: both mothers used either strategies, often mixing conventions within one situation:

(3) V. 1;7 M: *Varja chego vyp'jet?*
'What will Varja drink?'
M: *Chego ty vyp'jesh'?*
'What will you drink?'

or even in one utterance:

(4) K. 1;11 M: *Ne dam, a to mama sejchas nakazhet, nakazhu.*
'I won't give [it], or mummy will punish [you] soon, I will punish [you].'

The phenomenon of neutralising the person category in a child-directed speech by adults is very common and has been described in a great deal of works, see e.g. Chiat 1982, 1986, Durkin, Rutter and Tucker, 1982, Smoczyńska 1992. Nevertheless, it is difficult to give a simple answer to the question why such a habit exists.[9] Smoczyńska (1992) pays attention to the fact that parents – often unintentionally – want to make the process of understanding and acquiring a language easier to their children. Children are exposed to various forms, new lexemes, among which, pronouns may be perceived as difficult. If pronouns are substituted by more distinct first names or names of relatives (mummy, daddy etc.), a verb automatically must be expressed in 3:SG. It is required by the grammatical system. Hence, very often adults do not intentionally give up using personal forms of verbs, but they merely want to omit pronouns which, in their opinion, are difficult for the child to comprehend.

Secondly, it seems that adults often use 3:SG forms for ennobling, solemn purposes or in order to emphasise their role in a family, or – on the other hand – the child's importance. The phenomenon of bestowing a solemn rank on 3:SG could be observed in Kirill's recordings. For example, any time a new session begins, the mother, in order to trigger the conversation, addresses her son in a 3:SG form:

(5) K. 2;10 M: *Kirjusha budet rasskazyvat' mame knizhku?*
'Will Kirjusha tell mummy the book?'
K: *Plo (=pro) Chipollino.*
About Chipollino:SG:ACC
'About Chipollino.'

When she speaks as an authority, she also often uses more dignified, stronger form of 3:SG to refer to herself:

(6) K. 2;3 M: *Esli mama govorit, chto mozhno e~ti jagodki kushat', znachit ix mozhno rvat' i est', a esli nel'zja, to mozhno umeret'.*
'If the mother says you can eat those berries, it means you can pick them up and eat them, and if she says you cannot, it means you can die.'

Praising her son, she uses 3:SG form too:

(7) K. 2;3 M: *Kakoj Kirjusha bol'shoj mal'chik, sam pizhamku nadevaet.*

‘Oh, what a big boy Kirjusha is, he is putting his pyjamas on by himself.’

Differently, Varja’s mother treated ‘neutralised’ forms of 3:SG as one of the possible ways to communicate her child:

(8) V. 1;6 V: *Ain’ka (=Varen’ka) kapait (=kopaet).*
Varen’ka:SG:NOM dig:IPFV:PRES:SG:3
‘Varen’ka is digging.’
M: *Varen’ka kopaet?*
‘Is Varen’ka digging?’

Her utterances including verbs in 3:SG could rarely be interpreted as having a solemn implied meaning:

(9) V. 1;7 M: *Xoroshie devochki edjat kotletki i morkovki, da, i xodjat guljat’s mamami.*
‘Good girls eat cutlets and carrots, yes, and they go for walks with their mummies.’
M: *Vot Varja takaja xoroshaja devochka kak raz, kotoraja est kotletki i morkovki i xodit guljat’s mamoj.*
‘And Varja is such a good girl, who eats cutlets and carrots and goes for walks with her mummy.’

2.1. The frequency of personal forms in the input language

Percentages of personal forms used by either of the mothers in non-past tenses, calculated for the whole corpora, are presented in Table 3. The results in both corpora are very close: in singular and in plural. For the needs of this paper, however, only the distribution of singular forms is important: 1:SG and 2:SG as the standard ways of marking the speaker and the hearer and 3:SG, which may also be used in the same purpose, as it was already said above. Nevertheless, in order to illustrate the complete view of the development, the full distribution of all personal forms in the non-past tenses is presented in the table below:

Table 3. Percentage of personal forms in non-past tenses in the speech directed to children

	VARJA'S MOTHER	*KIRILL'S MOTHER*
SG:1	14	11
SG:2	29	32
SG:3	39	36
PL:1	12	11
PL:2	-	1
PL:3	6	10
TOTAL	100	100
N	**1818**	**7246**

The forms of 1:SG constitute 14% in Varja's mother and 11% in Kirill's. Forms of 2:SG appear much more frequently: in both cases they constitute about 1/3 of all non-past tense forms. 3:SG forms were the most frequent: in both corpora they constitute more than 1/3 of all the analysed forms.

3. ADDRESSER AND ADDRESSEE CATEGORY IN VARJA'S SPEECH

The frequencies of personal forms of non-past tenses in Varja's data are presented in Table 4:

Table 4. Percentage of personal forms in non-past tenses in Varja's data

	1;2	*1;3*	*1;4*	*1;5*	*VII 1;6-1;9*	*VIII 1;9-2;0*	*IX 2;0-2;3*	*X 2;3-2;6*	*XI 2;6-2;9*	*XII 2;9-3;0*	*whole corpus*
SG:1	-	8	5	9	20	38	40	50	39	36	**34**
SG:2	-	-	-	6	4	11	17	16	19	20	**12**
SG:3	(100)	68	62	66	61	40	32	26	32	27	**41**
PL:1	-	24	33	15	10	7	5	2	3	8	**7**
PL:2	-	-	-	-	-	-	-	-	-	1	**>1**
PL:3	-	-	-	3	5	4	7	6	7	9	**6**
TOTAL	100	100	100	100	100	100	100	100	100	100	100
N	**1**	**38**	**58**	**202**	**1159**	**1083**	**742**	**751**	**307**	**577**	**4918**

Let us first have a look at the appearances and raw frequencies of singular forms. First, at the age of 1;2, the 3:SG form appears (*pati '=plachet'* 'cry-IPFV:PRES:SG:3').: it was, undoubtedly, creatively used. By the end of VII quarter, the percentage of 3:SG forms was over 60%. In VIII quarter it dropped to 40%, and later on it fluctuated at the range of 30%.

The isolated form of 1:SG (*idu* 'go-IPFV:PRES:SG:1') appeared in the material from 1;3, but it was decidedly not a productive usage. In the following month I registered first productive uses: 1:SG forms were yet rare then: their frequency was about 5%. Their percentage was also low at 1;5, it was 9%, and afterwards, beginning from VII quarter it started growing. In VII quarter, those forms constituted 1/5 of all personal forms of non-past tenses spotted in Varja, and from VIII quarter – their share in the material was 40% (in X quarter, those forms constituted exactly half of all personal forms used by the girl).

At 1;5, with a frequency of about 6%, verbs in 2:SG appeared. In VII quarter their percentage was only 4%, and since VIII quarter, it started growing. In VIII quarter, it was 11% and later on it oscillated at the level of 16-20%.

The order of personal forms' occurrence in singular in Varja's corpus was thus the following: the girl first used verbs in 3:SG, afterwards in 1:SG and eventually in 2:SG. This order is, therefore, different from the one proposed by Chiat (1986).

3.1. 1:SG forms of non-past tenses in Varja's material

At the early stage of Varja's language development, up to 1;6, verbs in 1:SG occurred sporadically in the data, see Table 3. I assess that independent usages of this form occurred in the girl's data since 1;4: those were negated forms *ni atju (=ne xochu* 'do+not+want-IPFV:PRES:SG:1*),* and *budu (=ne budu* 'will+not-IPFV:FUT:SG:1*).* In the successive month, 1;5, 1:SG forms were built on the basis of different new stems:

(10) V. 1;5 V: *[v] magazin* *idu.*
shop:ACC:SG go:IPF:PRES:SG:1
'I'm going to the shop.'

(11) V. 1;5 V: *Abyk" (=jabloko)* *dezim (=derzhim),*
apple:SG:ACC hold:IPFV:PRES:PL:1
abyka (=jabloko) *dizju (=derzhu)* *i*
apple:SG:ACC hold:IPFV:PRES:SG:1 *and*
duguju (drugoe) *abyka (=jabloko)*
second:SG:F:ACC apple:SG:ACC
dizju (=derzhu).
hold:IPFV:PRES:SG:1
'We're holding an apple, I'm holding an apple
and I'm holding another apple.'

1:SG forms were for the most part used with reference to the speaker (hence, in compliance with the standard). At 1;5-1;7, I identified, however, a few non-standard usages of 1:SG forms, which namely, referred to the hearer. The description of this phenomenon will follow in section 3.2.1.

3.2. 2:SG forms of non-past tenses in Varja's material

2:SG forms of non-past tenses started to appear in Varja from 1;5: their frequency is presented in Table 3 above.

I have analysed the contexts in which 2:SG forms appeared and found out three types of usages: forms used with reference to the speaker (non-target like), forms used to mark the hearer (a standard deictic usage) and forms used in the phatic function. This last usage also conforms to the standard, especially for the spoken language, see Jakobson, 1989. The percentages of particular types are presented in Table 5:

Table 5. Percentage of 2:SG forms in non-past tense in Varja's data

SG:2 forms used	*1;2*	*1;3*	*1;4*	*1;5*	*VII 1;6-1;9*	*VIII 1;9-2;0*	*IX 2;0-2;3*	*X 2;3-2;6*	*XI 2;6-2;9*	*XII 2;9-3;0*
to mark the hearer	-	-	-	-	43	74	62	45	58	45
to mark the speaker	-	-	-	100	45	1	-	-	-	-
in phatic function	-	-	-	-	12	25	38	55	42	55
TOTAL	-	-	-	100	100	100	100	100	100	100
N	**0**	**0**	**0**	**12**	**49**	**117**	**127**	**117**	**57**	**117**

It stems from the data that first noted 2:SG forms, at 1;5, were used by the girl only with reference to the speaker. This phenomenon was still noted in VII quarter, however, on a smaller scale: less than half of registered 2:SG forms were used in such a way. The last such usages I spotted in the next quarter.

2:SG forms marking the addressee appeared in VII quarter: 43% of forms noted in this period were used in conformity to the standard as signifying the communiqué's receiver. In VIII quarter, this percentage increased to 74% and in IX quarter – it was more than 60%. In the following months, about half of registered 2:SG forms were references to the addressee and half – forms used in the phatic function. The latter appeared in the girl's corpus from VII quarter: then, they constituted 12%. Their frequency systematically increased later on until it reached the level of 50%.

3.2.1. 2:SG forms of non-past tenses in Varja's material: non-target like usages

The way of marking the communiqué's addresser through using 2:SG forms by Varja was clearly limited in time: I noted them in the period of 1;5-1;9, cf. Table 4. Here is one example from 1;7 just to illustrate the phenomenon:

(12) V. 1;7 M: *Chto ty delaesh'?*
what you do:IPFV:PRES:SG:2
'What are you doing?'
V: *Susais (=slushaesh') tupku (=trubku).*
listen:IPFV:PRES:SG:2 receiver:ACC:SG
'You are listening to the receiver.'

In the last example, dated at 1;9, Varja made a self-correction:

(13) V. 1;9 situation: Varja is touching a teddy-bear with her nose:
M: *Chto ty ego?*
'What are you [doing] to him?'
V: *Njuxaesh'.*
smell:IPFV:PRES:SG:2
'(you) are smelling.'
V: *Ja njuxaju.*
I smell:IPFV:PRES:SG:1
'I am smelling.'

As I have already mentioned (cf. section 3.1.), in the same age period, Varja several times (3 times at 1;5 and 1;6, twice at 1;7) used 1:SG forms erroneously in order to mark the hearer. To illustrate this, here is an example from 1;5:

(14) V. 1;5 M: *Chto ja derzhu?*
What I hold:IPFV:PRES:SG:1
'What am I holding?'
V: *Asjaku (=loshadku) dizju (=derzhu).*
horse:ACC:SG hold:IPFV:PRES:SG:1
'I am holding a horse'

This phenomenon, *person-reversal,*[10] is relatively seldom observed in children and is explained in various ways. E. Clark (1978, in Chiat 1982), investigating the English language, tried to explain the phenomenon in the following way. Some children construe personal pronouns as a kind of personal

names, as a sort of *fixed reference* to particular people. *I* would thus refer to adults and *you* – to children. Pronouns would not thus set roles in the discourse, i.e. They would not be used deictically.

This phenomenon is quite similarly interpreted by Oshima-Takane (1992). She also states that the personal pronouns are a kind of fixed reference to the particular people. The researcher sees, however, other reasons of such a treatment of pronouns by children. She analysed the material of the English speaking boy, David. According to Oshima-Takane, the source of the phenomenon lies in the fact that patterns of correct use of personal pronouns are not explicitly provided in the speech directed to children. The boy's mother, when speaking about herself, uses the pronoun *I*, and when addressing her son, she uses the pronoun *you*. When the boy speaks, he must reverse the pronouns, i.e. he has to know that pronouns are fixed to roles the participants of the conversation bear in the discourse. David was the only child and he spent most of his time with his mother or his nursemaid only. Therefore, he did not have many opportunities to observe how personal pronouns function in the discourse and he just did not know how to use them. The phenomenon of *person-reversal* is similarly described and explained by Vihman and Vija (this volume).

A completely different interpretation is proposed by Chiat (1982, 1986). Having analysed a lot of corpora of English speaking children, she found out that there was not a single child who used only reverse pronouns. Reversed-pronouns – if existing at all – appear in children's speech together with the target-like usages. The presence of both types of pronouns in children's speech shows that children are aware of the fact that roles of speaker and hearer constantly change during conversation. In Chiat's opinion, when children – as addressers – use a pronoun which would be used by the addressee, they take the addressee's perspective. Thereby, they extend the meaning of pronouns: on the one hand, they understand that pronouns set the participants of a discourse, on the other they are able to take different perspectives and they do that in a more varied way than the adults do. Chiat (1982) adds that in adults' language the personal pronouns happen to be multifunctional too. She exemplifies: *You may not smoke*. In this case, *you* may be interpreted either as deictic (*You may not smoke*) or as impersonal (*Smoking forbidden*). Perhaps, Chiat concludes, some children focus on the multifunctionality of pronouns and they distinguish this feature and extend it.

The above-mentioned attempts to explain the person-reversal trend do not seem, in my opinion, sufficient to interpret this phenomenon in Varja.The girl, despite the fact that she was the only child, had everyday opportunity to hear dialogues of her parents and grandparents. In the below-provided example, the grandmother addresses Varja's mother using the pronoun *you*, and afterwards she speaks about herself using the pronoun *I*:

(15) V. 1;7 situation: The women are talking about Varja:
B: *A chego ty segodnja reshila nosochki na nejo nadet'?*
'And why have you decided to put her the socks on today?'
M: *E~to doma.*
'It's only at home.'
B: *Ja uzh ispugalas'.*
'I have got frightened already.'

The girl could thus permanently observe in every-day conversation how the person category functions in the speech. Chiat's hypothesis may not be, also, deemed as the only possible interpretation of this phenomenon in the girl's case. In my opinion, there are strong correlations between Varja's being a *phrasal child* and the presence of utterances with reversed pronouns in her data.

(16) V. 1;5 situation: Varja doesn't want to dress:
I (=ne) xotit (=xochet) *tani (=shtany).*
not want:IPFV:PRES:SG:3 trousers:PL:ACC
'(she) doesn't want the trousers.'
I (=ne) xotit (=xochet) *botinki,*
not want:IPFV:PRES:SG:3 shoe:PL:ACC,
nimais' (=snimaesh')
take+off:IPFV:PRES:SG:2,
nimaj (=snimaj)
take+off:IPFV:IMPER:SG:2
'(she) doesn't want the shoes, you are taking off, take off.'

I would rather treat the utterances with erroneously used personal forms as one of the unanalysed 'formulae'. Such interpretation is supported with the fact the trend was restricted in time: when 'formulae' and erroneously used personal forms began to disappear from Varja's speech. Moreover, one of the Polish children, Kasia, whose data were gathered under professor Stefan Szuman's guidance in the 50s (cf. Smoczyńska 1998), was a *phrasal child* too and her data also included utterances with verbs in which the person was reversed (Smoczyńska 1992).

3.2.2. 2:SG forms of non-past tenses in Varja's material: standard usages

For the first time, Varja addressed the hearer using 2:SG forms at 1;6. Lexical analysis of all 2:SG forms used to mark the hearer showed that they are poorly diversified. Almost 1/4 of all usages of 2:SG forms of non-past

tenses referring to the hearer were represented by lexeme *xochesh'* 'want-IPFV:PRES:SG:2'. Nota bene, in VII quarter *xochesh'* constituted as many as 3/4 of all usages (!), in the following quarter – almost 50%. Further, the lexeme's frequency started decreasing – it balanced at the level of 6-11%. Varja's unconventional use of *xochesh'* is worth mentioning: asking for anything, the girl built sentences incorporating this lexeme:

(17) V. 1;11 *Batata, u menja ne*
granny, at I not
poluchaetsja, e~to bol'shaja
succeed:IPFV:PRES:SG:3 it:SG:NOM big
matrjoshka, xochesh' otkryt'?
matrjoshka:SG:NOM, want:IPFV:PRES:SG:2 open:INF
'Granny, I don't succeed, this is a big matrjoshka,
do you want to open it?'

I have obviously found *xochesh*' also in other phrases not expressing a request:

(18) V. 1;7 *Sobaka, xotis (=xochesh') aesyk (=oreshek)?*
dog:SG:VOC, want:IPFV:PRES:SG:2 nut:SG:ACC
'Dog, do you want a nut?'

As far as phatic usages of 2:SG forms are concerned, in Varja's data, the lexeme *vidish'* 'see-IPFV:PRES:SG:2' was predominant in this function, less frequently used was *znaesh*' 'know-IPFV:PRES:SG:2' – and it is in compliance with the standards of the spoken Russian language:

(19) V. 1;11 *Vot takie vot, prjamo, vidish', vot takie vot,*
tonen'kie, daj mne batatiny, batatskie konfetki,
znaesh', so stola.
'Just like this, simply, you see, just like this, thin,
give me granny's, granny's candies, you know,
from the table.'

3.3. 'Neutralised' 3:SG forms in Varja's data

It was difficult for me to decide how to present the frequency of 'neutralised' form of 3:SG non-past tenses. Total percentage characteristics of all 3:SG forms, such as in the case of 1:SG and 2:SG, where I have calculated the percentage of all forms in the entire corpus, says nothing about 'neutralised' forms. Hence, I decided to isolate 'neutralised' forms of 3:SG as a separate

category. The data are shown in Table 6. I have not stated the average result: it does not make any sense, because neutralisation of the person is a transitory phenomenon clearly limited in time.

Table 6. Percentage of 'neutralised' forms of 3:SG in Varja's data

	1;2	*1;3*	*1;4*	*1;5*	*VII 1;6-1;9*	*VIII 1;9-2;0*	*IX 2;0-2;3*	*X 2;3-2;6*	*XI 2;6-2;9*	*XII 2;9-3;0*
'neutralised' 3:SG forms	0	18	9	41	32	12	1	>1	0	**>1**
N of personal non-past tenses forms	1	38	58	201	1158	1083	742	751	307	**577**

At 1;3, the frequency of 'neutralised' 3:SG forms was 18% and afterwards it dropped to 9%. At 1;5, it culminated: the analysed forms constituted as many as 41% of personal forms of non-past tenses. Since VII quarter, their percentage in Varja's utterances increased: from 32% to 12% in VIII quarter. After that period the 'neutralised' forms of 3:SG appeared occasionally.

It stems from the data that 'neutralised' forms of 3:SG refer to the speaker, thus a little bit earlier than target-like 1:SG forms, began to appear in Varja's speech from the age of 1;3 onwards:

(20) V. 1;3 *Aika (=Varen'ka) titait' (=chitaet).*
Varen'ka:SG:NOM read:IPFV:PRES:SG:3
'Varen'ka is reading.'

I compared the frequencies of occurrences of forms referring to the speaker and hearer in Varja's and her mother's utterances respectively: the data are shown in graphs 1 and 2. Mother's utterances appear in the data covering the period between 1;2-1;3 very rarely: the data from this period are thus not regarded as representative.

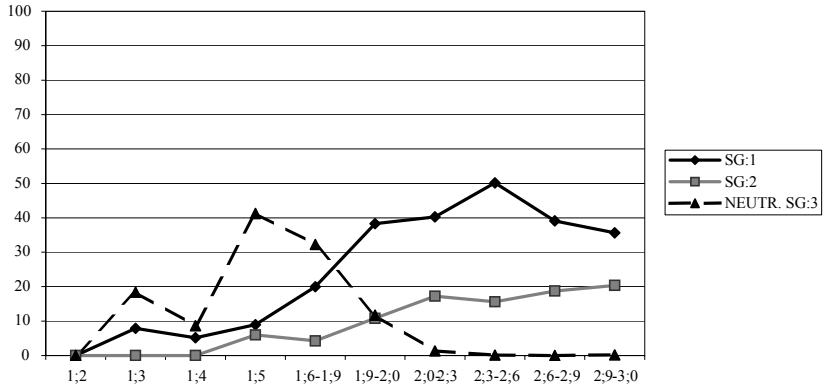


Figure 1. SG, 2:SG and 'neutralised' 3:SG forms in Varja's data

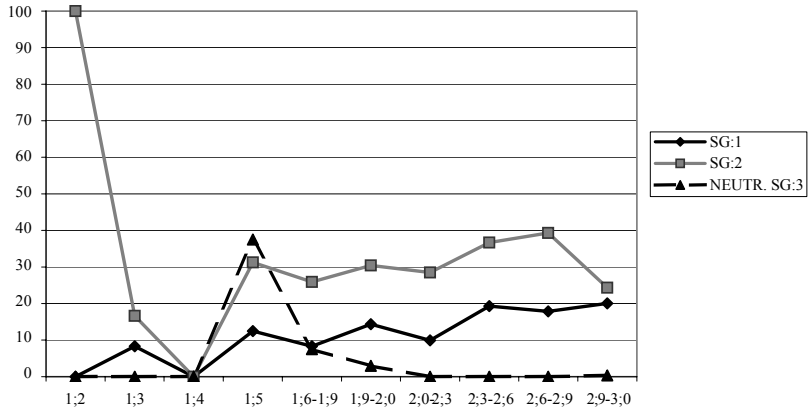


Figure 2. SG, 2:SG and 'neutralised' 3:SG forms in Varja's mother's data

The course of the line presenting the percentage of 'neutralised' forms of 3:SG is almost identical in both graphs, the daughter's and the mother's one. Culmination ensued at 1;5, later in VII and VIII quarters I noted a gradual decrease until such way of addressing the speaker and the hearer disappeared. Varja, like her mother, often changed the methods of marking the speaker and the hearer within one utterance:

(21) V 1;5 *I (=ne) budu* *isaat' (=risovat'), Aja (=Varja)*
not be:FUT:SG:1 draw:INF Varja:SG:NOM
i (=ne)xotit (=xochet).
not want:IPFV:PRES:SG:3
'I won't be drawing, Varja doesn't want.'

4. SPEAKER AND HEARER CATEGORY IN KIRILL'S DATA

Table 7 below contains the percentage analysis of personal forms of non-past tenses identified in Kirill's corpus:

Table 7. Percentage of personal forms in non-past tenses in Kirill's data

	1;8	*1;9*	*1;10*	*1;11*	*IX 2;0-2;3*	*X 2;3-2;6*	*XI 2;6-2;9*	*XII 2;9-3;0*	*whole corpus*
SG:1	-	(100)	75	15	45	42	42	39	**41**
SG:2	-	-	-	-	1	3	10	5	**5**
SG:3	-	-	17	65	41	40	36	39	**39**
PL:1	-	-	-	20	9	2	2	4	**5**
PL:2	-	-	-	-	-	-	-	-	**-**
PL:3	-	-	8	1	4	13	10	13	**10**
TOTAL	-	(100)	100	100	100	100	100	100	**100**
N	**0**	**4**	**12**	**82**	**570**	**612**	**1055**	**823**	**3156**

In the boy's speech, personal forms of non-past tenses occur from the age of 1;9 onwards. The first forms to appear are 1:SG. At 1;10 these forms constituted ¾ of all uses, at 1;11 – 15%. Since IX quarter, their percentage remains at more or less fixed level of about 40%.

At 1;10 I noted the first usages of 3:SG forms. They constituted then 17% of all registered forms. At 1;11 their percentage was high, it was more than 60%, further it remained at the level of 40%. As regards the singular number, verbs in 2:SG appeared at the latest: in IX quarter. In the entire analysed

material, their frequency came to a few percent: it oscillated around 10% only in X quarter.

The order of occurrence of personal forms in singular in the boy's material was thus the following: 1:SG-3:SG-2:SG. This schedule is close to Chiat's proposal.

4.1. 1:SG forms of non-past tenses in Kirill's material

The only independent forms of non-past tenses noted in the boy's corpus at 1;9 were verbs in 1:SG:

(22) K. 1;9 M: *Tvoja konfeta.*
'Your candy.'
K: *Kusju (=kushaju).*
eat:IPFV:PRES:SG:1
'I am eating.'

All registered 1:SG forms in the boy's data were used correctly.

4.2. 2:SG forms of non-past tenses in Kirill's material

2:SG forms of non-past tenses were rarely spotted in Kirill's material: their frequency did not exceed 5% in the entire corpus, see Table 6. However, the boy always used those forms in compliance with the standard: either to mark the hearer or in the phatic function, see Table 8:

Table 8. Percentage of 2:SG forms in non-past tense in Kirill's data

SG:2 forms used	*1;8*	*1;9*	*1;10*	*1;11*	*IX 2;0-2;3*	*X 2;3-2;6*	*XI 2;6-2;9*	*XII 2;9-3;0*
to mark the hearer	-	-	-	-	(100)	87	88	95
in phatic function	-	-	-	-	-	13	12	5
TOTAL	-			-	100	100	100	100
N	**0**	**0**	**0**	**0**	**3**	**16**	**101**	**38**

As it follows from the analysis, most forms of 2:SG used by Kirill referred to the hearer. In the example cited below, the boy easily made references to the speaker and the hearer by means of forms built on the basis of the same stem (*NB.* Kirill was the only child too!):

(23) K. 2;4 *Mne postroj domik!*
me:DAT build:PFV:IMPER:SG:2 house:SG:ACC
'Build me a house!'
M: *Sam postroj ty uzhe bol'shoj.*
'Build it yourself, you're a big [boy] already.'
Postroj sama,
build:PFV:IMPER:SG:2 yourself:SG:F:NOM,
ja ne umeju.
I not can:IPFV:PRES:SG:1
'Build [it] yourself, I can't.'
M: *Net, ty umeesh'.*
'No, you can.'
K: *A ja ne umeju,*
and I not can:IPFV:PRES:SG:1
a ty umeesh'!
and you can:IPFV:PRES:SG:2
'And I can't and you can!'

Kirill, unlike Varja, used phatic forms very rarely: they began to appear since X quarter: in X and XI with the frequency of 12-13%, in XII – they were only about 5%. Lexemes most commonly used by Kirill in this function included *vidish'* 'you see' and *znaesh'* 'you know':

(24) K. 2;7 *Takie den'gi, znaesh'*
such:PL:NOM money:PL:NOM know:IPFV:PRES:SG:2
kakie, tvoi.
which:PL:ACC your:PL:ACC
'Such money, you know which, yours.'

4.3. 'Neutralised' 3:SG forms in Kirill's data

The frequency of 'neutralised' forms of 3:SG in Kirill's material was very low. I noted only 16 such usages: the percentages are presented in Table 9.

Table 9. Percentage of 'neutralised' forms of 3:SG in Kirill's data

					IX	*X*	*XI*	*XII*
	1;8	*1;9*	*1;10*	*1;11*	*2;0-2;2*	*2;3-2;5*	*2;6-2;8*	*2;9-2;11*
'neutralised' 3:SG forms	**-**	**-**	**-**	**4**	**2**	**>1**	**>1**	**>1**
N of personal non-past tenses forms	**0**	**4**	**12**	**82**	**570**	**612**	**1055**	**823**

In Kirill, unlike in Varja, first references to the speaker were expressed in 1:SG form. 'Neutralised' verbs of 3:SG form appear from the age of 1;11 onwards:

(25) K. 1;11 K: *Djait' (=uezzhaet) Kiljusa (=Kirjusha).*
leave:IPFV:PRES:SG:3 Kirjusha:SG:NOM
'Kirjusha is leaving.'

The below-provided quotation dated at 2;0 is worth mentioning: the boy changed the convention of addressing the speaker within one situation:

(26) K. 2;0 K: *Podjom (=pojdjom) &iim – &po – &i pojdjom*
go:PFV:FUT:PL:1 go:PFV:FUT:PL:1
guljat'?
walk:IPFV:INF
'Will we go, will we go for a walk?'
K: *S"em tvojog (=tvorog) dja (=da)?*
eat:PFV:FUT:SG:1quark:SG:ACC, yes
'[when] I eat the quark, won't we?'
M: *Chto, Kirjusha?*
'What, Kirjusha?'
K: *My pojdjom guljat', dja (=da)?*
we go:PFV:FUT:PL:1 walk:IPFV:INF, yes
'Will we go for a walk [then], won't we?'
K: *S"est' Kijusja (=Kirjusha)*
eat:PFV:FUT:SG:3 Kirjusha:SG:NOM
tvoig (=tvorog), dja (=da)?
quark:SG:ACC, yes
'[when] Kirjusha eats the quark up, won't we?'

Next Kirill's utterance with the 'neutralised' form of 3:SG contains a solemn implied meaning (the boy was praised):

(27) K. 1;11 M: *Molodec.*
'Good boy.'
K: *Tatat' (=chitat') budet' (=budet').*
read:IPFV:FUT:SG:3
'[he] will read.'
M: *Chitaj, Kirjusha.*
'Read, Kirjusha.'

Figures 3 and 4 below present the frequency of forms referring to the speaker and the hearer in Kirill's and his mother's materials:

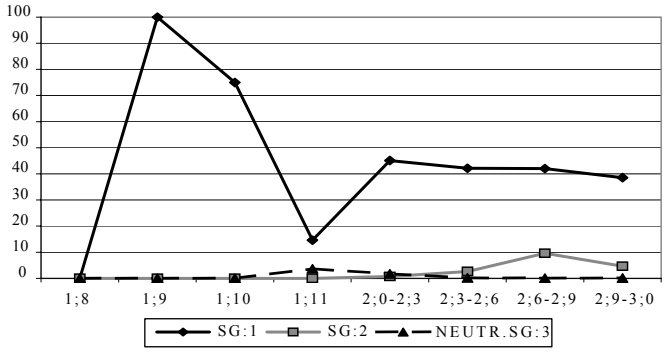


Figure 3. SG, 2:SG and 'neutralised' 3:SG forms in Kirill's data

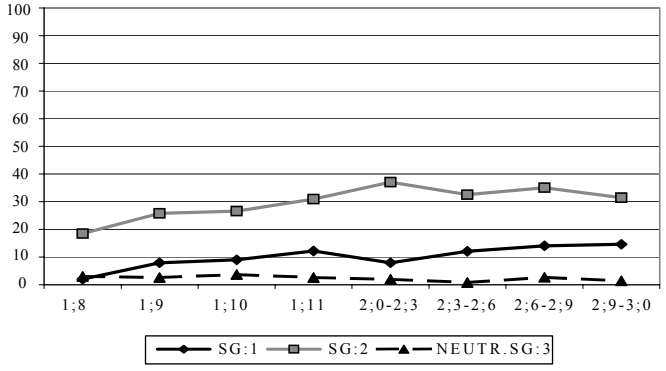


Figure 4. SG, 2:SG and 'neutralised' 3:SG forms in Kirill's mother's data

I see the same tendencies in Kirill and his mother as in Varja and her mother. First of all, 1:SG forms in boy were noted with a high frequency: since IX quarter, their percentage was above 40% of all personal forms of non-past tenses, 2:SG forms – with low frequency, about 5%. In his mother's

material – otherwise: the frequency of forms referring to the speaker was low, about 10%, whereas forms referring to the hearer were frequent: their percentage in the entire material was over 30%. The lines determining the frequency of 'neutralised' forms of 3:SG go almost identically in both graphs. The frequency of those forms in Kirill was a little above zero, and in his mother, it circulated around a few percent. What is important, the frequency of those forms in either corpora remained at the fixed level. Due to the low frequency of 'neutralised' forms in Kirill and his mother, I do not present qualitative analysis of them, i.e. reference 3/1 and 3/2. Their percentage is so low that it does not influence the general frequency of forms referring to the speaker and the hearer.

5. CONCLUSIONS

Finitive forms, i.e. those where person category is marked, appear in the Russian children's speech among the first acquired forms. It is a truism to state that not always and not all forms that appear at the earliest stage correspond with their functions and this statement can concern also personal forms. Nevertheless, after a very short period, when forms are used unproductively in the child language, when semantic distinctions can not be found, when forms are 'fossilized', there occurs a turning point: the child starts to build different forms based on the same root, it begins to be clearly seen how semantic of particular uses forms up and crystallizes.

Russian (like other Slavic languages) is a highly inflected language with a very rich morphology (both verbal and nominal). The fusion of several grammatical categories in one morpheme is a common phenomenon (e.g. in a verbal form pish-**u**:PRES:SG:1 'I am writing', the ending –u marks both: a person and a number). Paradoxically, the complex morphology makes the acquisition of Russian[11] easier for children: the children from the very early on can observe (and they use to) the necessity of continuous changes of word forms in particular contexts. So, high morphology in the input language accounts for quick formation of high morphology in the child language: we can thus talk about language-specific mechanisms.

The system of personal endings in non-past tenses – at the background of the whole grammatical system of Russian language – is relatively simply. There are no homonymic forms among the endings, the endings are clear and specific for each person. Typically (i.e. not language delayed) developing child should thus not have any problems in building the linguistic representation of personal distinctions.

In the above-presented analysis I described two case-studies. Both children had no difficulties with the mastering of marking personal oppositions, nevertheless both of them adopted completely different strategies at the earliest

stages of language development. Let me remind, that the girl preferred 'neutralised' forms of 3:SG for marking a speaker and a hearer. Only after a certain time, she switched onto canonical ways of marking the participants of a speech act, i.e. 1:SG form for speaker and 2:SG for hearer. Kirill from the very beginning used canonical forms of 1:SG and 2:SG, 'neutralised' forms of 3:SG were spotted in his data extremely rarely and their usage was strongly semantically marked. There arises a question, why the children adopted different strategies in acquiring a person category? I would suggest that both children ideally conformed to the pattern proposed by the environment (cf. Figures 1-4 above). In other words, they reflected their input. I would argue with the thesis, that 'neutralised' 3:SG forms are easier, simpler or clearer to children. I think children perceive both ways of marking the addresser and the addressee of a communiqué simply as alternative ways, not differentiating the more difficult and the easier one.

I would suggest that the order of particular personal forms' occurrences in both analysed children (let me remind that in Varja the order was the following: 3:SG-1:SG-2:SG and in Kirill: 1SG-3:SG-2:SG) is caused directly by the strategies the children adopted in the early phase of language development.

Finally, in every child, after the early stage of the language system formulation, when the frequency of particular forms is not yet fixed, there comes a period of stabilisation – the frequencies of certain forms are comparable both in children and in adults. The utterance structures of Varja and Kirill were comparable to the adults' material before the children were 2;6 (cf. Kiebzak-Mandera,1999). The percentages of different verbal forms (tense and aspect forms, moods etc.) were almost the same. Hence, the similar results could have been expected from the analysis of personal forms in children's speech. The percentages of forms referring to the addresser and addressee of the communiqué, however, range consequently in a different way in the children and their mothers. In Varja's material 1:SG forms of non-past tenses constituted 34% on average, in Kirill – 41%. In the speech directed to Varja and Kirill, these forms were 14% and 11% respectively. The percentage of 2:SG forms was 12% in the girl and only 5% in the boy. 2:SG forms constituted 29% of all personal forms of non-past tenses used by Varja's mother and 32% of the forms used by Kirill's mother.The low frequency of 2:SG forms in the children's material could suggest that this form is difficult to master. I studied how many 'neutralised' forms of 3:SG used by the children referred to the hearer. The results for Varja's material are presented in Table 10:

Table 10. Reference to the speaker and the hearer with 'neutralised' 3:SG forms in Varja's data (percentages)

	1;2	*1;3*	*1;4*	*1;5*	*VII 1;6-1;9*	*VIII 1;9-2;0*	*IX 2;0-2;3*	*X 2;3-2;6*	*XI 2;6-2;9*	*XII 2;9-3;0*
*'neutralised' forms 3/1**	-	93	95	98	73	90	100	(100)	-	(100)
*'neutralised' forms 3/2**	-	-	-	2	27	10	-	-	-	-
TOTAL	-	100	100	100	100	100	100	(100)	-	(100)
N	**-**	**7**	**5**	**83**	**373**	**127**	**10**	**1**	**-**	**1**

*The symbols come from Stojanowa-Trajkowa and Bojkowa (1987): '3/1' means that 3:SG form was used instead of 1:SG, '3/2' – 3:SG form was used instead of 2:SG form.

It stems from the analysis that Varja addressed the hearer with 'neutralised' forms quite rarely: at 1;5 about 2% of 'neutralised' forms of 3:SG applied to the addressee, in VII quarter – less than 1/3, and in VIII quarter – about 10%. In the subsequent period, the girl did not use 'neutralised' forms in this way: since IX quarter, the girl stopped using 'neutralised' forms at all. As regards Kirill, among 16 'neutralised' forms of 3:SG identified in his data, 5 referred to the hearer. Thus, the problem lies not in potential difficulties the mastering of 2:SG forms might have made: the children simply address the hearer rarely. Therefore, I can conclude that children and adults assume different roles in conversations and they consequently stick to them. Adults provoke children to speak – that is why there are a lot of 2:SG forms in their utterances. Children, having been provoked, give answers – in the broad meaning of this word. The quality of 'neutralised' 3:SG forms also speaks in favour of this hypothesis: most of these forms refer to the speaker. The results of the analysis might be different if the adults had not acted under a specific pressure, i.e. if they had not been aware, the material was being gathered. It often happens that parents want to fill up the running tape at all costs – for various reasons (because they feel obliged to fill the tape from the beginning to the very end, or they simply want to show off their child). Perhaps, that is why they provoke their children to speak so intensively. The fact that the frequencies of personal forms are spread so differently does not certify, by any means, to children's disadvantage. It rather proves their language skills and their capability to match the assumed or imposed roles.

DOROTA KIEBZAK-MANDERA
Jagiellonian University

6. NOTES

[1] The article is an amended version of one of the chapters of my unpublished doctoral thesis.
[2] In singular, only gender is marked, in plural, even this opposition is neutralised.
[3] The data were collected within *Crosslinguistic Project on Pre&Proto-morphology in the early language acquisition,* sponsored by the Austrian Academy of Sciences.
[4] Early stages of Varja's language development are described also in Kiebzak-Mandera, Smoczyńska and Protassova 1997.
[5] *Ibidem.*
[6] Boy's material is analysed in my unpublished doctoral thesis.
[7] A letter V stands for Varja, K for Kirill, M for their mothers and B for Varja's grandmother.
[8] It is worth noting that 'neutralised' 3:SG forms were used by the mothers almost only in the indicative: in the imperative this phenomenon was very rarely observed.
[9] However, in some families this is the only acceptable polite form in everyday language. Then, obviously, this dilemma does not arise.
[10] After Chiat (1982), the term *pronoun-reversal* has been broadly used by psycholinguists. In Russian, however, in non-past tenses, a personal ending attached to a verbal stem is the marker of person, a personal pronoun does not have to be used at all – that is why I decided to amend this term in the above-mentioned way.
[11] Cf. also the paper by Czinglar, Katičić, Köhler and Schaner-Wolles *Strategies in the L1-acquisition of predication: The copula construction in German and Croatian* (this book, pp. xx)

7. REFERENCES

Bates, Elisabeth., Philip. S. Dale, and Donna Thal. "Individual differences and their implications for theories of language development." In Paul Fletcher, Brain MacWhinney (eds.), *The Handbook of Child Language*, Oxford: Basil Blackwell, 1995: 96-151.

Charney, Rosalind. "Speech roles and the development of personal pronouns." *Journal of Child Language* 7 (1980): 509-528.

Chiat, Shulamuth. "If I were you and you were me: the analysis of pronouns in a pronoun-reversing child." *Journal of Child Language* 9 (1982): 359-379.

Chiat Shulamuth. "Personal pronouns." In Paul Fletcher and Michael Garman (eds.), *Language acquisition: studies in first language development.* Cambridge: Cambridge University Press, 1986: 339-355

Durkin, Kevin, D. R. Rutter, and H. Tucker. "Social interaction and language acquisition: Motherese help you." *First Language* 3 (1982): 107-120.

Grammatika. Fonetika. Fonologija. Udarenie. Intonacija. Slovoobrazovanie. Morfologija. vol. 1. N. J. Schvedova (ed.). Moskva: Izdatel'stvo Nauka, 1980.

Gvozdev, Aleksandr N. *Formirovanie u rebjonka grammaticheskogo stroja russkogo jazyka*. Moskva: Izdatel'stvo akademii pedagogicheskix nauk RSFSR, 1949.

Jakobson, Roman. "Poetyka w świetle językoznawstwa." In Maria Renata Mayenowa (ed.) *W poszukiwaniu istoty języka,* vol. 2. Warszawa: Państwowy Instytut Wydawniczy, 1989: 77-124.

Kiebzak-Mandera, Dorota. "Diary vs. tape-recorded data. Acquisition of Russian (and Polish) verbal morphology." 7th International Congress for the Study of Child Language, Istanbul, July 14-19, 1996.

Kiebzak-Mandera, Dorota, Magdalena Smoczyńska, and Ekaterina N. Protassova. "Acquisition of Russian verb morphology: the early stages." In Wolfgang U. Dressler (ed.), *Studies in of Pre-&Proto-morphology*. Wien, Verlag der Österreichischen Akademie der Wissenschaften, 1997: 101-114.

Kiebzak-Mandera, Dorota. Kształtowanie się systemu werbalnego u dzieci rosyjskojęzycznych: analiza porównawcza. Unpublished doctoral dissertation, 1999. Jagiellonian University, Kraków, Poland.

Kiebzak-Mandera, Dorota. "Formation of the verb system in Russian children." *Psychology of Language and Communication* 4 (2000): 27-46.

MacWhinney, Brian. *The CHILDES Project: Tools for analysing talk.* Hillsdale, NJ: Lawrence Erlbaum, 1991.

Laskowski, Roman. "Czasownik." In Renata Grzegorczykowa, Roman Laskowski, and Henryk Wróbel (eds.), *Gramatyka współczesnego języka polskiego*. Warszawa: Państwowe Wydawnictwo Naukowe, 1984: 225-269.

Oshima-Takane, Yuriko. "Analysis of pronominal errors: a case study." *Journal of Child Language* 19 (1992): 111-131.

Pine, Julien N. and Elena V. M. Lieven. "Reanalysing rote-learned phrases: Individual differences in the transition to multi-word speech." *Journal of Child Language* 20 (1993): 551-571.

Smoczyńska Magdalena. "Kategoria osoby w języku dziecka: odniesienie do nadawcy i odbiorcy." In Barbara Bokus and Maciej Haman (eds.), *Z badań nad kompetencją komunikacyjną dzieci.* Warszawa: Energeia, 1992: 205-228.

Smoczyńska Magdalena. "Krakowskie dane językowe w systemie CHILDES." In Magdalena Smoczyńska (ed.), *Studia z psychologii rozwojowej i psycholingwistyki. Tom poświęcony pamięci Marii Przetacznik-Gierowskiej.* Kraków: Universitas, 1998.

Stojanowa-Trajkowa, J. and S. Bojanowa. "Przyswajanie formalnych i pragmatycznych aspektów kategorii osoby przez małe dziecko." In Ida Kurcz, Grace Wales Shugar, and Barbara Bokus (eds.), *Wiedza a język,* vol. 2: *Język dziecka.* Wrocław, Warszawa, Kraków, Gdańsk, Łódź: Zakład Narodowy im. Ossolińskich, PAN, 1987: 71-80.

INDEX

STUDIES IN THEORETICAL PSYCHOLINGUISTICS

1. L. Solan: *Pronominal Reference.* Child Language and the Theory of Grammar. 1983 ISBN 90-277-1495-9
2. B. Lust (ed.): *Studies in the Acquisition of Anaphora.* Volume I: Defining the Constraints. 1986 ISBN 90-277-2121-1; Pb 90-277-2122-X
3. N. M. Hyams: *Language Acquisition and the Theory of Parameters.* 1986 ISBN 90-277-2218-8; Pb 90-277-2219-6
4. T. Roeper and E. Williams (eds.): *Parameter Setting.* 1987 ISBN 90-277-2315-X; Pb 90-277-2316-8
5. S. Flynn: *A Parameter-Setting Model of L2 Acquisition.* Experimental Studies in Anaphora. 1987 ISBN 90-277-2374-5; Pb 90-277-2375-3
6. B. Lust (ed.): *Studies in the Acquisition of Anaphora.* Volume II: Applying the Constraints. 1987 ISBN 1-55608-022-0; Pb 1-55608-023-9
7. G. N. Carlson and M. K. Tanenhaus (eds.): *Linguistic Structure in Language Processing.* 1989 ISBN 1-55608-074-3; Pb 1-55608-075-1
8. S. Flynn and W. O'Neil (eds.): *Linguistic Theory in Second Language Acquisition.* 1988 ISBN 1-55608-084-0; Pb 1-55608-085-9
9. R. J. Matthews and W. Demopoulos (eds.): *Learnability and Linguistic Theory.* 1989 ISBN 0-7923-0247-8; Pb 0-7923-0558-2
10. L. Frazier and J. de Villiers (eds.): *Language Processing and Language Acquisition.* 1990 ISBN 0-7923-0659-7; Pb 0-7923-0660-0
11. J.A. Padilla: *On the Definition of Binding Domains in Spanish.* Evidence from Child Language. 1990 ISBN 0-7923-0744-5
12. M. de Vincenzi: *Syntactic Parsing Strategies in Italian.* The Minimal Chain Principle. 1991 ISBN 0-7923-1274-0; Pb 0-7923-1275-9
13. D.C. Lillo-Martin: *Universal Grammar and American Sign Language.* Setting the Null Argument Parameters. 1991 ISBN 0-7923-1419-0
14. A.E. Pierce: *Language Acquisition and Syntactic Theory.* A Comparative Analysis of French and English Child Grammars. 1992 ISBN 0-7923-1553-7
15. H. Goodluck and M. Rochemont (eds.): *Island Constraints.* Theory, Acquisition and Processing. 1992 ISBN 0-7923-1689-4
16. J.M. Meisel (ed.): *The Acquisition of Verb Placement.* Functional Categories and V2 Phenomena in Language Acquisition. 1992 ISBN 0-7923-1906-0
17. E.C. Klein: *Toward Second Language Acquisition.* A Study of Null-Prep. 1993 ISBN 0-7923-2463-3

STUDIES IN THEORETICAL PSYCHOLINGUISTICS

18. J.L. Packard: *A Linguistic Investigation of Aphasic Chinese Speech.* 1993 ISBN 0-7923-2466-8
19. J. Archibald: *Language Learnability and L2 Phonology:* The Acquisition of Metrical Parameters. 1993 ISBN 0-7923-2486-2
20. M.W. Crocker: *Computational Psycholinguistics.* An Interdisciplinary Approach to the Study of Language. 1996 ISBN 0-7923-3802-2; Pb 0-7923-3806-5
21. J.D. Fodor and F. Ferreira (eds.): *Reanalysis in Sentence Processing.* 1998 ISBN 0-7923-5099-5
22. L. Frazier: *On Sentence Interpretation.* 1999 ISBN 0-7923-5508-3
23. S. Avrutin: *Development of the Syntax-Discourse Interface.* 1999 ISBN 0-7923-5936-4
24. B. Hemforth and L. Konieczny (eds.): *German Sentence Processing.* 2000 ISBN 0-7923-6104-0
25. M. De Vincenzi and V. Lombardo (eds.): *Cross-linguistic Perspectives on Language Processing.* 2000 ISBN 0-7923-6146-6
26. S.M. Powers and C. Hamann (eds.): *The Acquisition of Scrambling and Cliticization*. 2000 ISBN 0-7923-6249-7
27. M. Schönenberger: *Embedded V-to-C in child grammar: The acquisition of verb placement in Swiss German*. 2001 ISBN 0-7923-7086-4
28. M. Walsh Dickey: *The Processing of Tense*. Psycholinguistic Studies on the Interpretation of Tense and Temporal Relations. 2001 ISBN 1-4020-0184-3; Pb 1-4020-0185-1
29. C. Hamann: *From Syntax to Discourse*. Pronominal Clitics, Null Subjects and Infinitives in Child Language. 2002 ISBN 1-4020-0439-7; Pb 1-4020-0440-0
30. S. Nooteboom, F. Weerman and F. Wijnen (eds.): *Storage and Computation in the Language Faculty.* 2002 ISBN 1-4020-0526-1; Pb 1-4020-0527-X
31. A.T. Pérez-Leroux and J. Muñoz Liceras (eds.): *The Acquisition of Spanish Morphosyntax.* The L1 / L2 Connection. 2002 ISBN 1-4020-0974-7; Pb 1-4020-0975-5
32. H. Verkuyl, H. De Swart and A. Van Hout (eds.): *Perspectives on Aspect*. 2005 ISBN 1-4020-3230-7; Pb 1-4020-3231-5
33. N. Gagarina and I. Gülzow (eds.) *The Acquisition of Verbs and their Grammar: The Effect of Particular Languages.* 2006/2008 ISBN 978-1-4020-4334-5; Pb 978-1-4020-4336-9

STUDIES IN THEORETICAL PSYCHOLINGUISTICS

34. M. Bader and J. Bayer: *Case and Linking in Language Comprehension*. Evidence from German. 2006 ISBN 1-4020-4343-0