THE DEVELOPMENT OF ARGUMENT REPRESENTATION: A CROSSLINGUISTIC DISCOURSE-PRAGMATIC ANALYSIS OF ENGLISH AND JAPANESE CHILD LANGUAGE

A M Sonia Guerriero

Department of Psychology

McGill University

Montréal, Canada

August 2005

A thesis submitted to McGill University

in partial fulfilment of the requirements of the degree of Doctor of Philosophy.

Your file Votre référence
ISBN: 978-0-494-25160-7
Our file Notre référence
ISBN: 978-0-494-25160-7

CONTENTS

ABSTRACT

Children's learning of language-universal and language-specific principles of argument representation was the topic under investigation in the three studies comprising this thesis. Another objective was to investigate whether a discourse-pragmatic approach could be employed to explain children's patterns of argument omission and production, developmentally and crosslinguistically. To answer these questions, referential choice in the spontaneous language of monolingual English-speaking and monolingual Japanese-speaking children and their mothers was developmentally investigated whereby a sentence argument's morphological form (null, pronominal, lexical), referential status (given, new), and syntactic location (transitive subject, transitive object, intransitive subject) were systematically analysed. The first and second studies revealed that neither the English-speaking nor the Japanese-speaking children showed sensitivity to the referential distinction between given and new information early on in development (at 21 months of age). The English-speaking children mastered English-specific referential conventions between MLU 2.00 and 3.99 (between 24 and 32 months) and employed non-linguistic pragmatic correlates to supplement unconventional argument use from as early as MLU 1.00 (between 21 and 23 months). By contrast, the Japanese-speaking children showed unconventional referential choices as late as MLU 4.00 (between 33 and 36 months), as well as inconsistent use of non-linguistic pragmatic correlates. The third study revealed that, although language-specific differences were observed, neither group of children violated any of the four Preferred Argument Structure (PAS) constraints: The

children avoided using more than one new or lexical argument per transitive clause and avoided casting new or lexical arguments as transitive subjects. However, evidence of sensitivity to PAS strategies from early on in development was inconclusive because the children omitted most sentence arguments at the beginning of speech production. Finally, all three studies revealed that children's referential choices that were inconsistent with expected discourse-pragmatic principles reflected similar patterns observed in parental input. Altogether, this set of studies led to the following general conclusions regarding the learning of argument representation and distribution in syntax: (1) a discourse-pragmatic approach can explain language-universal features of argument omission and production in child language and (2) language-specific strategies are learned via parental input.

RÉSUMÉ

L'objectif principal de cette thèse, incluant trois études, était d'étudier l'apprentissage des principes d'argument de représentation du langage universel et du langage spécifique chez les enfants. Un autre objectif était de vérifier si une approche de discours pragmatique pouvait être employée pour expliquer les tendances des enfants en ce qui concerne les omissions et les productions d'arguments, en termes de développement et à travers les langues. Pour répondre à ces questions, le choix référentiel dans la langue spontanée d'enfants anglophones monolingues, d'enfants monolingues japonophones et de leurs mères a fait l'objet d'une recherche dans laquelle la forme morphologique (nulle, para nominale, lexicale) d'une phrase argumentative, ainsi que le statut référentiel (donnée, nouvelle) et l'emplacement syntaxique (sujet transitif, objet transitif, sujet intransitif) ont été analysés de façon systématique. Les deux premières études ont révélé que ni les anglophones, ni les japonophones ne montraient de sensibilité quant à la distinction référentielle entre de l'information donnée et de l'information nouvelle pendant leur développement précoce (à 21 mois). Les anglophones maîtrisaient les conventions référentielles spécifiques à l'anglais entre un MLU de 2.00 et de 3.99 (entre 24 et 32 mois), et employaient des corrélatifs pragmatiques pour compléter l'utilisation d'arguments non-conventionnels dès le MLU de 1.00 (entre 21 et 23 mois). En revanche, les enfants japonophones démontraient des choix référentiels non conventionnels à un MLU de 4.00 (entre 33 et 36 mois), ainsi qu'une utilisation inconsistante de corrélatifs pragmatiques non-linguistiques. La troisième étude a révélé que bien que des différences

spécifiques à la langue avaient été observées, aucun des groupes d'enfants n'enfreignaient les quatre contraintes de la « Preferred Argument Structure », ou structure d'argument préférentielle (PAS) : les enfants évitaient d'utiliser plus d'un argument nouveau ou lexical par clause transitive ainsi que d'employer des arguments nouveaux ou lexicaux comme sujets transitifs. Cependant, les signes de la sensibilité des enfants aux stratégies de la PAS pendant le développement précoce n'étaient pas conclusifs car les enfants omettaient la plupart des arguments dans la phrase au début de la production de la parole. Enfin, les trois études ont démontré que les choix référentiels des enfants qui étaient incohérents avec les principes de discours pragmatique attendus reflétaient des tendances similaires observées dans la contribution parentale. Cette série d'études a mené aux conclusions générales suivantes concernant l'apprentissage de représentation d'argument et de distribution de syntaxe : (1) une approche pragmatique du discours peut expliquer les particularités de l'omission et de production d'argument chez l'enfant en termes de langage universel et (2) les stratégies spécifiques à la langue s'acquièrent via la contribution parentale.

CONTRIBUTIONS OF AUTHORS

This thesis consists of three manuscripts. I am first author on all manuscripts and my graduate supervisor, Yuriko Oshima-Takane, is a co-author. A third author, Yoko Kuriyama, appears on the first and third manuscripts (i.e., Chapters 2 and 4).

All three papers were researched, written, and revised by myself. The research design for the three studies was developed by myself in collaboration with Yuriko Oshima-Takane and the coding scheme was developed by myself, with guidance from Yuriko Oshima-Takane. I was responsible for analysing and interpreting the data and Yuriko Oshima-Takane provided editorial feedback and statistical advice. She also provided the English data (videotapes and transcripts), and I performed the reliability for the transcripts (verifying the accuracy of the transcripts with the videotapes) and the coding of the data. I trained and supervised the native Japanese speakers for the transcription, reliability, and coding of the Japanese transcripts. They are personally acknowledged at the beginning of each manuscript. The Japanese data reported in the second paper (i.e., Chapter 3) were obtained from the CHILDES database. Yoko Kuriyama provided the Japanese data (videotapes and a first draft of the transcripts) reported in the first and third studies and, as such, is listed as a third author.

Portions of each study have been presented at various conferences: the Society for Research in Child Development, Minneapolis, April 2001 and Tampa, April 2003, the joint conference of the IXth International Congress for the Study of Child Language and the 23rd annual Symposium on Research in Child Language Disorders, Wisconsin, July

2002, and the 17th biennial meeting of the International Society for the Study of Behavioural Development, Ottawa, August 2002 (Guerriero, Oshima-Takane, Cooper, & Kuriyama, 2001; 2002a; Guerriero, Oshima-Takane, & Ono, 2002b; 2003). Portions of the first and third studies appear in the *Proceedings of the Annual Boston University Conference on Language Development* (Guerriero, Cooper, Oshima-Takane, & Kuriyama, 2001). The first manuscript has been submitted to the Journal of Child Language and is currently in revision. Based on the reviewers' feedback, the first and second studies will be combined into one paper. The third manuscript is in preparation for submission.

ACKNOWLEDGEMENTS

First and foremost, I would like to extend my sincere appreciation to my graduate advisor, Yuriko Oshima-Takane. I thank her for all her invaluable support throughout my graduate career, constructive and timely feedback on my thesis, and the many enlightening discussions on theoretical issues. She has been a great mentor and instrumental in making me the researcher I am today.

I would also like to extend my gratitude to my graduate committee members, Fred Genesee, Tom Shultz, and Rachel Mayberry, for their feedback on the three manuscripts reported in this thesis, and in particular to Fred Genesee for his encouragement and support.

Analysis of the Japanese data would not have been possible without the help of my native Japanese-speaking coders: Amy Cooper, Yuhko Kayama, and Alyssa Ono. They served not only as research assistants, but as friends as well. Many thanks to Yuhko for letting me audit her "Introduction to Japanese" course.

I could not have pursued graduate studies without the endless love and support of my family and in-laws. Words cannot express how grateful and fortunate I feel to have them in my life. Finally, my deepest and most heartfelt thanks go to my husband, Walter. I thank him for his love and encouragement and for reminding me every day of what is really important in life. This thesis is for him.

CHAPTER 1

General Introduction

Young children at early stages of language development tend to omit subject or object arguments from their spontaneous productions. It is a prominent and universal feature of early child language and has been characterized as the "null argument phenomenon." The following are some examples:[1]

(1) a. Ø want go get it.

b. Ø helping Mommy.

c. Mummy get Ø.

d. Man taking Ø.

e. Ø put on Ø.

The sentences above, collected from naturalistic child data, are lacking either the subject argument (1a and 1b), object argument (1c and 1d), or both (1e). Such argument omissions generally persist until about 3;0 or 4;0 years of age (Allen, 1997) and have been observed crosslinguistically.

The three studies presented in this thesis are, for one thing, an investigation of the null argument phenomenon, and secondly, an investigation of how children come to learn the grammatical use of subject and object arguments in spoken sentences, as per the specifications of the target language.

---

[1] Examples are cited from L. Bloom (1970) and Radford (1990). The Ø symbol represents the location of an omitted argument.

The investigation of children's null arguments has been a subject of debate for some time.[2] Traditionally, explanations of the underlying or mitigating causes have taken one of two theoretical approaches, namely, competence- or performance-based approaches. In recent years, a third approach, the discourse-pragmatic approach, has shown some gains in popularity. The approach I take in this thesis is a discourse-pragmatic one, but the competence- and performance-based approaches are also briefly reviewed and evaluated in this introductory chapter. Before examining each approach, I first review the patterns of children's argument omissions that have been reported in the literature.

Patterns of Children's Argument Omissions

Universality of Omission

One of the distinguishing features of the null argument phenomenon is that it occurs crosslinguistically, whether children are learning an overt or a null argument language. Null argument languages are those languages which allow argument omission. In this group of languages, speakers omit arguments as a grammatical option, usually when the discourse context or morphosyntactic rules of the language allow the omission. Some languages allow either subject or object arguments to be omitted, such as Chinese, Korean, Japanese, as well as American Sign Language; others allow only the subject argument to be omitted, such as Italian, Spanish, and Portuguese. In some null subject languages, such as Spanish and Italian for instance, the subject argument, typically a

[2] I use the terms "null" and "omitted" interchangeably to indicate when a verbal argument is not overtly produced in speech.

pronominal, can be freely omitted because its referent is easily understood via the inflectional markings on the verb. Thus, subject omission is grammatical due to the morphosyntactic rules of the language. In other null argument languages, such as Chinese, Korean, and Japanese, the syntax carries no overt information with respect to the omitted argument. The referent is typically understood via the discourse and pragmatic context (Wang, Lillo-Martin, Best, & Levitt, 1992). Even though omission in null argument languages is a grammatical option, cases of unacceptable omission might occur if omission were improperly used where an overt argument would have been more appropriate. Null argument languages are thus different from languages, such as Danish, French, and English, which do not permit argument omission. In overt (non-null) argument languages, argument omission is not an option; speakers must produce as overt all subject and object arguments or the utterance is deemed ungrammatical.

However, regardless of whether the target language is a null or an overt argument one, all children progress through a stage whereby arguments are omitted. The null argument phenomenon has been observed and investigated in overt argument languages such as English (L. Bloom, 1970; P. Bloom, 1990; Gerken, 1991; Hyams, 1986; Ingham, 1992; O'Grady, Peters, & Masterson, 1989; Radford, 1990; Valian, 1991; Wang, Lillo-Martin, Best, & Levitt, 1992), German (Clahsen, 1986; 1991; Hamann, 1995; 1996; Weissenborn, 1992), Dutch (de Haan & Tuijnman, 1988), Danish (Hamann & Plunkett, 1998), Swedish (Strömqvist & Ragnarsdóttir, 2000), Icelandic (Sigurjónsdóttir, 1999; Strömqvist & Ragnarsdóttir, 2000), and French (Weissenborn, 1992); in null argument languages such as Japanese (Hirakawa, 1993; Ingham, 1992; Mazuka, Lust, Wakayama, & Snyder, 1986; Nakayama, 1994), Chinese (Wang, Lillo-Martin, Best, & Levitt, 1992),

Inuktitut (Allen, 1997; 1999; 2000), Korean (Clancy, 1993; 1997), and American Sign Language (Lillo-Martin, 1986b; Lillo-Martin, 1994), in null subject languages such as Italian (Hyams, 1986; Rizzi, 1994; Valian, 1991), Brazilian Portuguese (Valian & Eisenberg, 1996), European Portuguese (Valian & Eisenberg, 1996), Hebrew (Braine, 1974; Uziel-Karl & Berman, 2000), and Spanish (Austin, Blume, Parkinson, Núñez del Prado, & Lust, 1997b), as well as in a null object language such as Sesotho (Demuth, Machobane, & Moloi, 2000).

Although omitted arguments have been observed in both null and overt argument languages, one feature distinguishes between the two: Children learning overt argument languages eventually begin to supply overt arguments more and more consistently, whereas children learning null argument languages continue to omit them. There is thus a difference in developmental patterns between the two language typologies and this is further discussed in the next section

Inconsistency of omission

Whether children are learning a null or an overt argument language, argument omission is inconsistent, as sentences containing omitted arguments co-occur simultaneously with sentences containing overt arguments within the same linguistic stage. For example, Andrew, observed by Braine (1963), produced sentences such as "plug in" and "Andrew plug in" between 2;0 and 2;3. This inconsistency has led some to refer to the null argument stage as one during which children "optionally" omit (or use) sentence arguments (L. Bloom, 1970; P. Bloom, 1993; de Haan & Tuijnman, 1988; Hyams, 1986; Ingham, 1992; Jaeggli & Hyams, 1987; Mazuka, Lust, Wakayama, &

Snyder, 1986; Nakayama, 1994; O'Grady, Peters, & Masterson, 1989; Radford, 1990; Valian, 1990; Valian, 1991; Valian & Eisenberg, 1996). Although the phenomenon of child argument omission has been investigated for some time now, it has only infrequently been investigated developmentally. The few studies reporting developmental data are summarized in Table 1. The table gives the authors, language, design, and a summary of the main findings for each study.

For those languages which do not permit argument omission, inconsistent use gradually gives way to more consistent use as children begin producing overt arguments on a more frequent basis. Omission patterns differ between subject and object arguments. In contrast to object omission, which is reported to be more consistent throughout, subject omission is more rampant at earlier MLU stages than at later ones (e.g., Valian, 1991). The developmental pattern of omission (or use) is highly variable. For instance, subject omission in two children learning English (Eve and Adam) was reported to fluctuate during the omission phase until it eventually tapered off (P. Bloom, 1993; Hyams & Wexler, 1993). An unswerving linear decline is not usually observed.

In contrast to overt argument languages, the same decline in argument omission is not observed with children learning null argument languages. Moreover, children's patterns resemble adult patterns. Nakayama (1994) analysed argument omission in three Japanese-speaking children and reported that subject and object omission rates did not change much from 1;0 to 2;6 and that the children showed rates similar to adult Japanese speakers. Table 1 presents findings from null argument, as well as null subject, languages that have been investigated developmentally.

In sum, when compared to children learning overt argument languages, children learning null argument or null subject languages do not show a sharp decline in their use of null arguments. They continue to omit arguments across MLU periods. On the other hand, argument omissions for children learning overt argument languages gradually decline over time until they largely disappear, at which point overt arguments become predominant. Regardless of language typology, however, there is no such period of absolute omission at early stages of language development. Omitted arguments co-occur simultaneously with overt arguments and, as such, there is no abrupt or sudden shift from omission to overt usage.

Null Subject/Null Object Asymmetry

Finally, researchers report a disparate distribution between subject and object omission and a tendency for subject arguments to be omitted with greater frequency than object arguments. This, too, has been observed in overt as well as null argument languages and has come to be known as the "null subject/null object asymmetry." For example, Hirakawa (1993), who analysed the spontaneous speech of one Japanese-speaking girl between 2;0 and 3;1, obtained means of 89.1% versus 38.9% for subject versus object omission, respectively (also Nakayama, 1994). Evidence of an asymmetry has been observed in other null argument languages, such as Inuktitut (Allen, 1997), Chinese (Wang, Lillo-Martin, Best, & Levitt, 1992), and Hebrew (Uziel-Karl & Berman, 2000). Studies looking at overt argument languages report a similar null subject/null object asymmetry, also with a greater incidence of subject omission. P. Bloom (1990), who investigated the null subject/null object asymmetry with three English-speaking

children aged between 1;6 and 2;7, reported a 9% rate of object omission versus a 55% rate of subject omission. A null subject/null object asymmetry in English child language has similarly been reported by L. Bloom (1970), Bloom, Miller, and Hood (1975), Gerken (1991), Menyuk (1969), Valian (1991), and Wang, Lillo-Martin, Best, and Levitt (1992).

In sum, three facts are established: The first is that regardless of language typology, there is an asymmetry between subject and object omissions, and the second is that the asymmetry is due to greater subject omission. Third, the asymmetry occurs in both null and overt argument languages. Although a null subject/null object asymmetry has been widely reported, not all agree to its universality (e.g., Mazuka, Lust, Wakayama, & Snyder, 1986; Radford, 1990).

Goals of the Thesis:

Issues in Child Argument Omission and Production

The purpose of this thesis was to investigate the development of argument representation in child language. Understanding children's argument representation involves investigating not only argument omission, but also how children come to learn the (grammatical) use of overt subject and object arguments in spoken language. One of my primary goals, thus, was investigating whether child argument omission is random or whether there is some sort of systematic pattern which can predict when, or under what circumstances, arguments are omitted. Another of my goals was to investigate argument omission and production using a discourse-pragmatic perspective. I chose a discourse-pragmatic approach because the competence- and performance-based approaches were

found to be inadequate in accounting for the patterns of argument omission described previously. I describe the discourse-pragmatic approach in some detail in this chapter, after evaluating the competence- and performance-based approaches. Here, I outline the issues in child argument omission that led me to select the discourse-pragmatic theory as the most adequate. The first three issues have already been discussed: (1) the universality of omission, (2) the inconsistency of omission (and gradual shift towards overt production for learners of overt argument languages), and (3) the null subject/null object asymmetry. A good theory is one that accounts for all three of these patterns.

The flip side to investigating which arguments are chosen for omission is to analyse which arguments are chosen for production. Thus, a fourth requisite of a good theory is one that can also explain the realization of arguments. Given that young children are limited to producing two, sometimes three, words at a time for any one utterance, the question is which arguments will they choose to produce (e.g., Allen, 2000; Greenfield & Smith, 1976)? An adequate theory is one that can account for the arguments children choose to omit as well as predict which arguments they choose to realize as overt.

A fifth issue concerns the developmental nature of language: The same theory that explains the omission and production of arguments in early grammar must also explain the omission and production that occurs across development and in later adult grammar. Particularly for those children learning null argument languages, a good theory must explain whether arguments are omitted with the same motivation across developmental periods or whether the motivation underlying argument omission in later language differs from that underlying omission in early language. This is an important issue since a theory

of acquisition should be consistent with adult grammar. A related issue has to do with children acquiring overt argument languages and how they eventually learn that their language obligatorily requires overt arguments. In other words, how do these children "unlearn" the omission? A good theory should be able to explain the unlearning; and a theory which can explain both the production and omission of arguments within the same theoretical framework, not just at early stages, but in adult grammar as well, is one that presumes the continuity of processes throughout development.

A final issue has been alluded to before. It concerns the obvious typological differences between null and overt argument languages: Are two explanations necessary, one to account for null argument languages, the other for overt argument ones? A more powerful, convincing theory of child argument representation is a universal one; one which can explain and predict the arguments that are chosen for omission and production and their underlying motivation, regardless of the typological differences among languages.

With these issues in mind, I briefly review performance- and competence-based theories of argument omission and explain why they are inadequate in explaining child argument representation. Likewise, the same issues are used to explain why the discourse-pragmatic approach was employed in the present set of studies.

Competence-Based Approaches

Competence-based approaches, also referred to as grammatical or linguistic approaches, have been the most frequently examined theories. According to these approaches, grammatical phenomena are explained using a Principles and Parameters

framework, a framework based on the assumption that humans are born with a Universal Grammar (UG) containing a set of universal linguistic principles true of all languages. It is assumed that children are born knowing the basic linguistic structure of each phrase (such as the Noun Phrase, Verb Phrase, Adverbial Phrase, and so on), but that the language-specific ordering of the phrasal elements must be learned. Since children initially lack the proper orderings, they are said to represent grammars differently from their adult counterparts and, as such, lack appropriate adult "competence." Children's grammars are not fundamentally different from adult grammars, only less mature. Later, based on positive evidence regularly provided from parental input, children's grammars mature or develop into the appropriate target grammars (Haegeman, 1991).

A plethora of competence-based accounts of the null argument phenomenon have been put forth with respect to the specific linguistic processes that induce omission, such as optional infinitive accounts (e.g., Guilfoyle, 1984; Hoekstra & Hyams, 1996; Hyams, 1996; 2001; O'Grady, Peters, & Masterson, 1989; Sano & Hyams, 1994; Schütze & Wexler, 1996; Wexler, 1994), topic-drop accounts (e.g., Bromberg & Wexler, 1995; de Haan & Tuijnman, 1988; Hamann, 1995; Hyams & Wexler, 1993), truncation accounts (e.g., Haegeman, 1996; Rizzi, 1994), parameter-setting accounts (e.g., P. Bloom, 1990; 1993; Clahsen, 1991; Hyams, 1986; 1992; Hyams & Wexler, 1993; Jaeggli & Hyams, 1987; Lillo-Martin, 1986a; 1994; Mazuka, Lust, Wakayama, & Snyder, 1986; Rizzi, 1994; Valian, 1990; 1994; Wang, Lillo-Martin, Best, & Levitt, 1992; Weissenborn, 1992), and structure-building accounts (e.g., Radford, 1990; 1996).

Parameter-setting accounts, in particular, have been the most popular. The main tenet of such accounts is the assumption that whether or not languages allow argument

omission as a grammatical option is the result of a variation on a parametric setting. Languages that allow omission, such as Italian, which allows subject omission, are said to have the "Null Subject Parameter" switched on. To account for argument omission in child language, parameter setting accounts posit that the default setting for the Null Subject Parameter is [+null subject]. In other words, all children begin with the assumption that their language is a null subject one. Children who deviate from the adult target, such as children whose target language is overt argument, are seen as having "mis-set" the parameter. Gradually, mis-set parameters are re-set into ones more in harmony with the target adult language, made on the basis of positive evidence provided in parental input (Hyams, 1992; Hyams & Wexler, 1993; Jaeggli & Hyams, 1987). There have been many variations of parameter-setting approaches (e.g., the Principle Branching Direction Parameter of Mazuka, Lust, Wakayama, & Snyder (1986) and the Discourse-Oriented and Null-Pronoun Parameters of Wang, Lillo-Martin, Best, & Levitt (1992)), but regardless of the specifics, the mechanism inducing the omission is always the same: The setting (or mis-setting) of a parameter, thus allowing argument omission as a "grammatical" option.

There are disagreements as to the appropriate default setting. For example, P. Bloom (1990; 1993) suggested that the default value is that argument omission is *not* allowed, and as such, children assume that overt arguments are obligatory. Then, those children who hear omission in their input, such as children exposed to null argument or null subject languages, make a parametric "switch" to the appropriate setting. Valian (1990; 1994), on the other hand, proposed that all parameters are initially un-set. There are no default settings; rather, children have both values of a parameter available, as

equally possible alternatives, as well as a set of hypothesis-testing procedures by which to evaluate and weigh input evidence. By detecting and analysing the patterns and regularities children hear in their input, one value of the parameter will begin to weigh more than the other.

Evaluation of Competence-Based Approaches

Approaches of a competence-based nature have the advantage of explaining linguistic phenomena in a systematic manner by invoking (what are assumed to be) universal grammatical principles. Proponents of these approaches claim that the advantage of their proposals over performance and discourse-pragmatic proposals lies in their explanatory power in that they predict the development of several seemingly unrelated linguistic structures through the operation of one mechanism, such as a switch in parameter setting. For instance, it has been proposed that the co-occurrence of the end of the null subject phase and the appearance of various properties of the inflectional system (e.g., Clahsen, 1986; 1991; Guilfoyle, 1984; Hyams & Wexler, 1993; Jaeggli & Hyams, 1987; Lillo-Martin, 1986a; Pierce, 1994) or the co-occurrence of infinitives and subject omission (e.g., Hamann & Plunkett, 1998; Hoekstra & Hyams, 1996; Hyams, 1996; Radford, 1990; Rizzi, 1994; Sigurjónsdóttir, 1999) are developmentally related. As the child's grammar develops into the target adult grammar, each newly attained linguistic stage generates a wide range of effects which are thought to be related (de Haan & Tuijnman, 1988; Hyams, 1986; 1992; Jaeggli & Hyams, 1987; Mazuka, Lust, Wakayama, & Snyder, 1986). Thus, the correlation between subject omission and various grammatical milestones allows approaches of a grammatical nature to account for a range

of phenomena in a principled fashion. Given the rapid rate at which children learn language, the simultaneous acquisition of several grammatical features is an obvious advantage.

However, when tested empirically, the proposed relations between subject omission and various properties of the inflectional system or the relation between subject omission and infinitival verb forms have not been confirmed (e.g., P. Bloom, 1993; Hamann & Plunkett, 1998; Ingham, 1992; O'Grady, Peters, & Masterson, 1989; Valian, 1991; Weissenborn, 1992). A relation between a parameter switch and the emergence of a grammatical property will always be difficult, if not impossible, to establish. This is because the development of linguistic milestones is never abrupt, but rather gradual. For instance, a switch from one parameter setting to another predicts a sudden, abrupt shift from a [+null subject] to a [–null subject] grammar. In reality, however, the development from null to overt argument use is gradual, with overt arguments being used more and more consistently over time (P. Bloom, 1993; Hyams & Wexler, 1993; Valian, 1991; Wang, Lillo-Martin, Best, & Levitt, 1992; Weissenborn, 1992). In addition, the simultaneous occurrence of both overt and omitted arguments, especially within the same linguistic stage, is not predicted by a linguistic theory designed to explain either one or the other behaviour. Others have also noted these complications of competence-based theories of acquisition (e.g., Austin, Blume, Parkinson, Núñez del Prado, & Lust, 1997b; P. Bloom, 1993; Ingham, 1992; O'Grady, Peters, & Masterson, 1989). Thus, competence-based accounts, by nature, are discontinuous theories. Such theories, however, unnecessarily complicate matters because, not only are different mechanisms necessary,

as opposed to just one, but by consequence such theories would also need to postulate how children "unlearn" their earlier grammars.

Another complication for competence-based theories has to do with explaining the null subject/null object asymmetry. Most accounts do not typically make any predictions with respect to object omission and claim such omissions to be a result of performance errors. Others have proposed that object omission is carried out by a different mechanism, thus resulting in a disparity between subject and object omission (Hyams, 1992; Jaeggli & Hyams, 1987). However when the predictions made by such mechanisms have been tested directly, they have not been supported (Hirakawa, 1993; Wang, Lillo-Martin, Best, & Levitt, 1992).

It is generally hypothesized that parental input is the driving force behind children's changing grammars, and this is an obvious theoretical advantage for competence-based theories since the role of input in language learning is clearly necessary (e.g., Choi & Gopnik, 1995; Miyata, Oshima-Takane, & Nisisawa, 2004; Naigles & Hoff-Ginsberg, 1995; Naigles & Hoff-Ginsberg, 1998; Tardif, Shatz, & Naigles, 1997; Theakston, Lieven, Pine, & Rowland, 2001). The role of input for competence-based theories is essential because they predict that it is a specific "trigger" in children's input that is responsible for inducing a grammatical change. For example, under a parametric account, children learning English will use the lack of null argument sentences in their input as evidence that their grammar is a [–null subject] one, thus inducing a parametric switch (e.g., Hyams, 1992). But postulating a *specific* trigger is also a complication because input is generally misleading. For example, children learning English will hear a significant number of imperatives (i.e., sentences that do not contain

overt subjects) in their early years; thus, their input will provide evidence for both null and overt subject sentences (see also Valian, 1990).

Performance-Based Approaches

Performance-based approaches claim that because young children's memory, attention, or processing abilities are less mature or less developed than those of proficient adult speakers, the "speaking" process exhausts children's limited resources resulting in performance errors during speech production (Bloom, 1970; Bloom, 1990; Bloom, 1993; Gerken, 1991; Pinker, 1984; Valian, 1991; Valian & Eisenberg, 1996).[3] As such, children's argument omissions are thought to arise at the level of speech production, not underlying competence. In this sense, performance-based proposals are like competence-based accounts, in that they are nativist in nature and assume universal and innate syntactic knowledge, or at least knowledge of universal syntactic principles and acquisition of the target competence from a very early age; as early as 2;0 (Bloom, 1993; Valian, 1991). Performance-based approaches also adopt a linguistic framework, such as the Principles and Parameters framework, in describing the mechanism(s) involved. Some parameter-setting accounts postulating a performance, rather than a competence, deficit include Mazuka, Lust, Wakayama, and Snyder's Principal Branching Direction Parameter (1986), P. Bloom's proposal that [–null subject] is the default setting (1990; 1993), and Valian's dual-value solution (1990; 1994). Non-parametric performance-

[3] I use the terms "performance deficit," "performance limitation," "processing deficit," and "processing limitation" interchangeably with identical meanings.

based accounts have also been proposed (see L. Bloom's (1970) reduction transformation account and Gerken's (1991) metrical account).

The exact nature of children's performance limitation(s) is still an open question. Valian (1991) and Valian and Eisenberg (1996) suggest that children have memory constraints due to the smaller capacity of their working memories. Whether the memory limitation is due to children's inferior ability to form chunks or to their inexperience with information processing itself (i.e., organizing, monitoring, and integrating information) is not known. As children develop and gain linguistic experience, speaking becomes more automatic and, thus, less cognitively demanding. The less deliberate the speaking process becomes, the more processing space becomes available, thus increasing working memory capacity and reducing processing limitations. For very young children, limited memory resources will force them to economize in production by performing as few computations as possible, resulting in omitted arguments from surface structure. L. Bloom (1970) argues that limitations in sentence production are not limitations in memory per se, but rather are constraints on the number of syntactic operations that may be performed within one sentence or constraints on the complexity of grammatical relationships. Thus, argument omissions in surface structure reflect the child's inability to carry the full sentence load of the underlying representation. Pinker (1984) suggests that a "processing bottleneck" occurs somewhere between communicative intention and the actual utterance. A young child can coordinate only a fixed number of lexical items at each stage between message conception and speech production because the child's processing mechanisms are limited in capacity. Thus, a bottleneck arises. The limitation is "not in

the child's rule system itself, but in the procedures that utilize the rules in speech production" (Pinker, 1984, p. 161).

When children's (limited or immature) processing resources are taxed, one way for them to economize their resources is to omit certain (obligatory) constituents from overt production. This allows children to produce sentences that are less complex, and thus shorter in length, than actually intended, generally resulting in omission of the subject (L. Bloom, 1970; P. Bloom, 1990; Mazuka, Lust, Wakayama, & Snyder, 1986; Valian, 1991; Valian & Eisenberg, 1996). Subject arguments are particularly affected because planning a sentence is thought to be more effortful at the beginning than at the end, causing the processing load to be higher at the beginning of a sentence than at the end (e.g., P. Bloom, 1990; Valian, 1991). Due to a more taxing processing load at the beginning of sentences, subjects tend to be omitted more frequently than objects, thus resulting in a null subject/null object asymmetry. P. Bloom (1990; 1993) reported that the "heaviest" subjects were associated with the shortest verb phrases (also Hyams & Wexler, 1993; Valian, 1991), indicating that lexical subjects required more processing resources than pronominal subjects, which in turn required more processing resources than omitted subjects.

More evidence in support of the claim that children suffer from production limitations that constrain syntactic complexity include: (1) a length effect whereby sentences with omitted subjects are longer than sentences with lexical (overt) subjects (L. Bloom, 1970; Bloom, Miller, & Hood, 1975; P. Bloom, 1990; Brown & Fraser, 1963; Ervin, 1964; Feldman, Goldin-Meadow, & Gleitman, 1978; Gerken, 1991; Valian, 1991; Valian & Eisenberg, 1996), (2) an increase in the use of overt subjects as development

progresses (P. Bloom, 1993; Hyams & Wexler, 1993; Kim, 1997; Valian, 1991; Valian & Eisenberg, 1996; Wang, Lillo-Martin, Best, & Levitt, 1992; Weissenborn, 1992), (3) an increase in the use of pure transitive verbs (verbs which obligatorily require an object such as *find, catch, take*) as development progresses (Theakston, Lieven, Pine, & Rowland, 2001; Valian, 1991), and (4) an increase in the use of overt objects with mixed transitive verbs (verbs such as *eat, smoke, read* for which an object is optional) as development progresses (Theakston, Lieven, Pine, & Rowland, 2001; Valian, 1991). Finally, it has been noted that, during the null argument period, subject arguments are sometimes not completely omitted, but merely reduced (L. Bloom, 1970). The reduction might be a partial verbalization of the subject argument and, as such, suggests that children at this stage might understand that a subject is required but are having difficulty in producing it, due perhaps to processing limitations (P. Bloom, 1990; Mazuka, Lust, Wakayama, & Snyder, 1986).

Evaluation of Performance-Based Approaches

One advantage of performance-based proposals is that the inconsistency of omission during the null argument phase (i.e., the co-occurrence of both omitted and overt arguments) is accounted for: Children are capable of producing all the required arguments, since they are presumed to have adult-like underlying grammatical competence, but because they suffer from processing limitations, they are sometimes inconsistent in their use of subject and/or object arguments in surface structure. Moreover, as children's linguistic abilities become more sophisticated, a performance account predicts a gradual decrease in ungrammatical argument omission. In this way,

there is no need to posit a mechanism for the "unlearning" of argument omission in the case of children learning overt argument languages.

P. Bloom's (1990; 1993; also Valian, 1991) prediction in support of the null subject/null object asymmetry, however, runs into complications. The hypothesis that lexical subjects are "heavier" than pronominal or omitted subjects was meant in regards to the phonological length of the subject noun phrase (e.g., *the little brown puppy* being longer and thus heavier than *he*). However, this concept is not consistent with linguistic theory. In linguistic theory, a pronominal argument (e.g., *he*) is linked to an antecedent in previous discourse (e.g., *the little brown puppy*); thus, pronominal subjects should presumably be "heavier" than lexical subjects. This is because the speaker would need to employ processing resources to not only make the cognitive link between pronoun and antecedent, but also to keep a mental record of what items were referred to in prior discourse. The same is true for the use of omitted arguments in null argument languages; the referents of these are linked to antecedents in prior discourse. Lexical arguments, on the other hand, are not generally linked to antecedents in prior discourse. Thus, contrary to a processing deficit account of the null subject/null object asymmetry, omitted and pronominal arguments should exert more, not less, of a processing load than lexical arguments.

Also problematic is that there has been a failure in finding solid evidence of a processing limitation. The predictions have either not been upheld or been found to be inconclusive. For example, a developmental increase in overt subject use is not observed crosslinguistically. Investigations of argument omission with children learning null argument languages do not report an increase in the children's overt subject use, but

rather a steady use of null arguments throughout or a stabilization of omission rates by the second year of life (Allen, 2000; Hirakawa, 1993; Kim, 2000; Nakayama, 1994). In fact, these studies report that children show similar rates of omission to adult speakers of those languages (Allen & Schröder, 2003; Austin, Blume, Parkinson, Núñez del Prado, & Lust, 1997b; Hirakawa, 1993; Kim, 2000; Valian, 1991; Wang, Lillo-Martin, Best, & Levitt, 1992). The lack of a corresponding increase in overt subject use with children learning null argument languages casts doubt on whether a performance deficit might actually be responsible for children's argument omissions. It seems unlikely that only children learning overt argument languages omit arguments due to insufficient processing resources. If performance limitations are inducing children's argument omissions, such limitations are predicted to affect the linguistic behaviour of all children, regardless of the type of language they are learning. Moreover, a performance account can be a viable explanation of children's argument omission only if a significant difference in omission rates is found to exist between children and adults. Finally, supporting evidence of children's adult-like grammatical competence has similarly proven elusive (L. Bloom, 1970; P. Bloom, 1990; 1993; Pinker, 1984; Valian, 1991; Valian & Eisenberg, 1996). For instance, recent studies seem to indicate that young children are incapable of discriminating between transitive (pure and mixed) and intransitive verbs (Theakston, Lieven, Pine, & Rowland, 2001; also Tomasello, 1992a), thus indicating immature linguistic competence.

In sum, there is no doubt that children surely suffer from processing limitations and that some sort of processing deficit is most likely involved in young children's early language. But that this processing deficit alone is responsible for how children choose to

represent arguments, whether to represent them as omitted or overt, is difficult to substantiate.

Discourse-Pragmatic Approaches

The basic tenet underlying a discourse-pragmatic analysis of argument representation is that pragmatic factors govern argument selection during discourse.[4] Discourse-pragmatic factors are thought to govern both the morphological form of arguments, such as whether an argument is represented as an omitted, pronominal, or lexical form, as well as the argument's distribution in syntax, such as whether the argument appears in subject or object position. Since this is the approach I employ in this thesis, I review the relevant theoretical background in some detail in the sections that follow.

The Mental Representation of Discourse Referents

Chafe (1987; 1994; 1996) suggested that the mental concepts representing discourse referents (i.e., referents mentioned during conversational discourse) reside in one of three "degrees of activation" in a person's mind, whether speaker or listener: (1) active, (2) semi-active, or (3) inactive. Each degree of activation is essentially a state of consciousness and a particular concept may be in only one of these three states at any particular time. A concept that is in the "active" state is said to be in the person's focus of

[4] I use the terms "referential form," "referential choice," "morphological form," "morphological choice," "argument form," "argument choice," "argument selection," "argument representation," and so on, interchangeably to mean the type of referring expression used by the speaker to represent an argument in syntax, whether the argument appears as a lexical, pronominal, or null form.

consciousness at the present moment. These are concepts that have been mentioned recently – or recently "activated" – in the discourse and are thus in the person's working memory. A concept remains active by virtue of its repeated mention in the discourse. A "semi-active" concept is said to reside in the person's peripheral consciousness. It is not being directly focussed on at the present moment, but is in the person's background awareness. It was focussed on (i.e., was active) at an earlier point in the discourse but has since receded from the fully active state. Such a concept has thus lapsed into a semi-active state during a period when it failed to be refreshed through repeated mention. Due to the limited resources of working memory, concepts that are not continually focussed on become semi-active. Semi-active concepts do not immediately become inactive but remain in a semi-active state for some time. As semi-active, concepts may once again become active via re-introduction into the discourse conversation. Concepts may also become semi-active via their association with a schema (Chafe, 1987). A schema is a cluster of interrelated concepts. When a concept is evoked during discourse, all its associated concepts enter the semi-active state. For instance, activation of a "school" concept will bring into semi-activation all school-associated concepts, such as "teacher," "classmates," "classroom," "desk," "books," "exams," and so on. Concepts associated with the physical environment of the discourse situation, such as all the items in the particular setting in which a conversation is taking place, are thought to also be peripherally semi-active, even though these are never directly focussed on. Finally, an "inactive" concept is one that is neither focally active, nor peripherally semi-active. Inactive concepts may be in a person's long-term memory or might never have entered consciousness before, in which case they would represent newly learned information.

The different activation states of active, semi-active, and inactive are relevant to a discourse-pragmatic theory of argument representation because these states reflect not only the status of concepts in a person's consciousness (i.e., memory), but also how concepts are represented in discourse. A concept that the speaker assumes is not known to the listener, and thus is inactive in the listener's consciousness, will be represented as "new" information to the discourse taking place. A concept assumed to be known by or shared with the listener, thus in the listener's active consciousness, will be represented as "old" information. In recent years, the term "given" has been used to designate information in active focus. Concepts are said to be represented as "accessible" information if the speaker assumes that they are semi-active in the listener's consciousness.[5] Stated another way, new information is characterized as that which is newly introduced (i.e., newly activated) in the discourse, given information as that which is already mentioned (i.e., already active) in the discourse, and accessible information as that which is activated from a previously semi-active state. Thus, the exchange of information between speaker and listener reflects the continuous movement of information in and out of consciousness. Chafe (1994) characterizes this dynamic process of information movement as the "flow of information in discourse."

---

[5] There is debate as to whether the informational status of discourse referents is speaker- versus listener-oriented. According to researchers such as Chafe (1994; 1996) and Prince (1981), speakers determine the informational status of discourse referents based on their assumptions of the status of such information in the mind of the listener. Others, such as Fisher and Tokura (1995) and Bard et al. (2000), on the other hand, claim that speakers do not take into account any specific information that may be known by the listener. Instead, speakers use their known knowledge of the discourse context in determining a discourse referent's informational status, not the listeners'.

Others have proposed slightly different definitions for what constitute given and new information. For instance, Clark and Haviland (1977) wrote that sentences convey two kinds of information: (1) Information that the speaker considers "given" is information believed to be already known and accepted as true by the listener, and (2) information that the speaker considers "new" is information that the listener does not yet know. Prince (1981) further develops the notion of given information by suggesting that "givenness" is characterized into three types: (1) predictability (concepts the speaker assumes the listener can predict or recover), (2) saliency (concepts the speaker assumes are in the listener's consciousness), and (3) shared knowledge (concepts the speaker assumes that the hearer already knows, assumes, or can infer).

The Cognitive Cost of Discourse Referents

Discourse-pragmatic theories inherently include the effect of performance factors on speech production. This is because the flow of information in discourse, or more accurately, the distinction between givenness, accessibility, and newness, in essence reflects what Chafe (1994; 1996) termed "activation cost." Due to the limited processing resources of working memory, a certain "cognitive cost" is associated with each mention of a concept in discourse, depending on whether the concept's earlier state was active, semi-active, or inactive. Given information is the least costly since it is already active. Such information has already been mentioned in the discourse and its active status in the mind of the listener does not require much mental effort in order for the referent to be identified. Accessible information is more costly than given information, and new information is the most costly of all. This is because more mental effort is involved in

converting a concept from the inactive to the active state. Previously inactive concepts are recalled from long-term memory and, as such, require greater cognitive effort than active or semi-active concepts. Schema-associated concepts in the semi-active state are more easily accessible than fully inactive concepts.

The Expression of Discourse Referents

A concept's activation cost essentially influences how it is verbally expressed in discourse. Chafe (1987; 1994; 1996) observed that the speaker's choice of referring expression and a referent's state of activation show a correlational relationship. Given concepts are typically expressed with a pronominal form or omitted altogether in the form of null arguments, if the language permits such an option. Given information is typically spoken with attenuated pronunciation, such as weaker stress (also Nooteboom & Kruyt, 1987). Certain concepts are, for cognitive reasons, always in the active state and thus always verbalized as given information. These include first- and second-person referents (i.e., the speaker and listener) because these are always present in the conversational context and because during a conversation each is conscious of the other as well as of herself or himself. Other concepts that are always in the active state include features of the universe, such as the sun and moon, body parts, kinship terms, such as "Mommy" and "Daddy," and proper names (also Givón, 1983). On occasion, newly introduced information might also be expressed as given (i.e., with a pronominal or null form) due to its saliency in the discourse context, such as the appearance of a third-person referent into the physical context, thereby putting it into the focal consciousness of the speaker and listener. Thus, givenness may be established either linguistically or

extralinguistically. Accessible information tends to be expressed in the same way as new information. Both of these are typically expressed as accented "lexicalized" forms, such as with a noun or a full noun phrase (also Nooteboom & Kruyt, 1987). Depending on the discourse situation, referents that are given might sometimes need to be accented, such as when a speaker wishes to contrast one referent with another, and are thus lexicalized into a full NP. At other times, a given referent might be lexicalized, but with weak, as opposed to strong, stress, if the speaker judges that the use of a pronominal or null form will cause ambiguity between two equally possible referents currently active in the mind of the listener (also Terken & Hirschberg, 1994). Generally, given information tends to be omitted or pronominalized, depending on which language-specific grammatical option is available in the target language; accessible and new information tend to be lexicalized. A predictable pattern is thus established between referential choice and discourse-pragmatic factors.

Others have made similar proposals. For instance, Givón (1983) suggested the "Topic Accessibility" scale; Ariel (1990; 1996) proposed an "Accessibility Marking Scale;" and Gundel, Hedberg, and Zacharski (1993) proposed a "Givenness Hierarchy." Although each proposal is different in its own right, they all have in common the objective to explain the representation of arguments in discourse.

The Distribution of Referring Expressions in Syntax

Discourse-pragmatic factors govern not only an argument's morphological form, as discussed above, but also how arguments are distributed in syntax. To fully explain how argument choice relates to syntactic distribution under a discourse-pragmatic

perspective, I briefly review some theoretical principles underlying information theory (e.g., Greenfield & Smith, 1976; Lambrecht, 1994).

The communicative function of an utterance (reflected linguistically as a sentence) is to transmit information from speaker to listener. Due to the limited capacity and processing resources of working memory, one pragmatic principle is that speakers should convey only those ideas that are informative and relevant to the discourse taking place. However, given the vast amount of information available in both the linguistic and extralinguistic discourse contexts, the question arises as to how the speaker decides which elements are informative and relevant. The decision is typically made based on the principle of "informativeness," which has to do with the degree of uncertainty surrounding a referent's identity in the situational context. Uncertainty exists where the situational context presents a novel or changing element or where several possible alternatives exist. The novel or changing element or the element among possible alternatives which reduces uncertainty the most is considered to be the most informative. Elements that are constant, unchanged, or old are perceived to be the most certain, and thus the least informative (Greenfield, 1979; Greenfield & Smith, 1976; Greenfield & Zukow, 1978). The principle of informativeness predicts that the discourse conversation should centre around those referents that are the least certain, and thus the most informative.

The certainty-uncertainty distinction essentially reflects the perceptual-cognitive distinction between the psychological concepts of presupposition and assertion in natural language discourse (Greenfield, 1979; Greenfield & Smith, 1976). For communication to be successful, the speaker and listener must share presupposed knowledge (Greenfield &

Zukow, 1978), so the speaker makes assumptions regarding what knowledge is presupposed on the part of the listener. Presupposed knowledge consists of not only the information communicated in the prior discourse, but also the extralinguistic information that is part of the situational context, as well as the unverbalized information that the speaker and listener share as implicit knowledge of the world; essentially, information that is already "activated" in the discourse. The distinction between presupposition and assertion thus relates to the certainty-uncertainty distinction: Certain elements are presupposed; uncertain elements are asserted. In this sense, a presupposed referent represents "uninformative" information, whereas an asserted one represents "informative" information (Baker & Greenfield, 1988; Bates, 1976; Greenfield & Smith, 1976; Greenfield & Zukow, 1978).

During conversational discourse, presupposed information, which is assumed rather than stated, tends to be omitted or represented with a pronominalized form. Assertions, on the other hand, cannot be taken for granted and are thus expressed with a lexicalized form, such as a noun phrase (Chafe, 1987; 1994; 1996; Lambrecht, 1994). Thus, presupposed information typically represents given information in discourse, whereas asserted information typically represents new information. In this sense, the notions of presupposition and assertion reflect the activation cost (i.e., mental representation) of discourse concepts and, as such, regulate how information is linguistically represented in discourse.

In terms of syntax, the information conveyed by a proposition (syntactically a sentence) is comprised of a combination of presupposed and asserted information, whereby newly asserted information is related to information that is already given, or

taken for granted, in the discourse. The typical discourse strategy used by adult speakers is to present one concept as a starting point and then add information about it. This strategy has traditionally been described as a topic-comment ordering strategy. Syntactically, the topic-comment discourse strategy is manifested as a subject-predicate structure, where the "topic," typically the grammatical subject, encodes given (presupposed) information and the "comment," typically the predicate, encodes new (asserted) information (Baker & Greenfield, 1988; Bates, 1976; Givón, 1983; Greenfield, 1979; Greenfield & Smith, 1976; Greenfield & Zukow, 1978). Vygotsky (1962), for instance, wrote about the "psychological subject," the element which is already known, and the "psychological predicate," the added information about the subject. Chafe (1987; 1994; 1996) proposed a "light subject constraint" (also called a "light starting point constraint") whereby speakers tend to choose subjects expressing referents that are already active (i.e., given) in the discourse situation because grammatical subjects act as "starting points" to which new information is then attached. This linguistic formatting strategy ensures that newly activated information is linked, or related, to given and presupposed information (Du Bois, 1987; Lambrecht, 1994). In this way, the linguistic manifestation of subject and predicate corresponds to the psychological notions of presupposition and assertion.

Discourse-Pragmatic Strategies in Adult Speech

Chafe (1987; 1994; 1996) analysed spontaneous and narrative conversations from adult speakers of English and Seneca (a Native American overt argument language spoken on reservations in western New York State) and reported that the starting point of

a clause typically contained a given referent, sometimes an accessible referent, but rarely a new referent; hence the "light subject constraint." Du Bois (1987; also 2003) analysed narrative discourse data from adult speakers of Sakapultek (a null argument Mayan language spoken in highland Guatemala) and reported that certain patterns of information and grammar tended to recur consistently, which he interpreted as a statistically "Preferred Argument Structure" (PAS). Chafe's and Du Bois's proposals are essentially the same; the difference being that Du Bois distinguished between transitive and intransitive clauses and differentiated subject arguments with respect to whether they were used with transitive or intransitive verbs. Du Bois's findings revealed that new arguments tended to be cast either as subjects of intransitive verbs or objects of transitive verbs, but rarely as subjects of transitive verbs, giving rise to the "Given A Constraint."[6] Similarly, lexical arguments tended to be cast either as subjects of intransitive verbs or objects of transitive verbs, but rarely as subjects of transitive verbs, giving rise to the "Non-Lexical A Constraint." These findings were interpreted as grammatical and pragmatic role constraints whereby speakers show a tendency to avoid casting lexical or new (i.e., "heavy") arguments in subject position of transitive verbs.

Du Bois (1987) also reported that clauses tended to contain either zero or just one lexical argument, but rarely two lexical arguments, giving rise to the "One Lexical Argument Constraint." Similarly, clauses tended to contain either zero or one new argument, but rarely two new arguments, which he termed the "One New Argument Constraint." These were interpreted as grammatical and pragmatic quantity constraints

---

[6] Du Bois (1987) used the following notation: "S" = subject argument of an intransitive verb; "A" = subject argument of a transitive verb; "O" = object argument of a transitive verb.

whereby speakers show a tendency to avoid more than one new or lexical argument per clause. Chafe (1987; 1994; 1996) similarly reported that each sentential clause rarely contained more than one new piece of information (in terms of activation cost), which he termed the "one new idea constraint." According to Chafe, a single clause (what he refers to as an "intonation unit") is constrained in the amount of new information that it can convey to the discourse, and that, due to the light subject constraint, the new information will most likely not be carried by the subject. Because an intonation unit typically consists of just one clause, the new information will most likely be carried in the predicate. In this way, Chafe's two constraints are related to Du Bois's four constraints. Studies with adult speakers of languages of varying typologies (e.g., Hebrew, English, German, Brazilian Portuguese, French, Spanish, Japanese, among others) indicate that such patterning is exhibited not only with narrative discourse, but with spontaneous and text data as well, reflecting what might be universal discourse-pragmatic patterning strategies (Ashby & Bentivoglio, 1993; Du Bois, 1987; Kärkkäinen, 1996; Matsumoto, 2000).

The rare occurrence of more than two lexical or two new arguments within one clause suggests that there exists an upper cognitive limit on the amount of information that can be fully active in the mind at any one time, likely related to the limitations of working memory. As such, these constraints speak to the processing limitations inherently contained in a discourse-pragmatic account of language production, which, according to Chafe (1987; 1994; 1996), likely result from the cognitive basis of a linguistic clause. Assuming that new information does in fact exert more cognitive effort as the information is converted from an inactive to an activated state, it is no surprise then

that new, and by consequence lexical, arguments are more difficult to process and thus more subject to cognitive restrictions.

Discourse-Pragmatic Strategies in Child Speech

Greenfield and Smith (1976) conducted a year-long study of the spontaneous speech of two children who were at the peak of the one-word stage. They reported that the one word which was verbalized was typically the one carrying the most uncertainty, where uncertainty, as defined by information theory, was due to the existence of several possible alternatives. For instance, agents were found to be infrequently expressed as isolated single-word utterances. This is because the children typically did not express elements that could be taken for granted given the situational context, and under normal circumstances the identity of an agent can usually be taken for granted. An agent was expressed only if its identity could not be assumed, if there was a conflict over agency, or if there was an actual change of agent. In other words, only under conditions of uncertainty, where alternative agents were available in the situation, was an agent expressed. Thus, even from a child's point of view, those elements which can be assumed are not stated, whereas those which cannot be taken for granted are verbalized. Children continue to use the principle of informativeness in choosing which elements to express even at the two-word stage (Baker & Greenfield, 1988; Bates, 1976; Greenfield & Zukow, 1978). These findings indicate that the principle of informativeness can be invoked to explain which element is selected for expression from as early as the one-word stage, as children seem to be sensitive to pragmatic principles of discourse from the beginning of speech production.

In regards to more linguistically sophisticated children, Allen (1997; 2000) and Clancy (1993; 1997) analysed referential choice in the spontaneous language data of four children acquiring Inuktitut and two children acquiring Korean, respectively. Both are null argument languages which allow omission of either subject or object arguments. The Inuktitut-speaking children were observed longitudinally for a period of nine months, the youngest child aged 2;0; the Korean-speaking children were observed for a period of one year, the youngest child aged 1;8. Overall, the results revealed that, for both languages, the children tended to realize as overt those arguments representing informative referents, whereas arguments representing uninformative referents tended to be omitted. Informativeness was measured by analysing features indicative of discourse, or pragmatic, prominence. A referent was characterized as pragmatically prominent, and thus informative in discourse, if it was (a) queried or a response to a query, (b) contrasted with another referent, (c) absent from the context, (d) new (i.e., mentioned for the first time), (e) ambiguous with another referent in the context, (f) ambiguous with another referent in the discourse, (g) inanimate, or (h) third person. A referent classified as containing one or more of these pragmatic features would be less certain, in the sense of information theory, and would thus require the child to be as informative as possible when choosing a referring expression. For languages such as Inuktitut and Korean, which allow argument omission, being informative would mean verbalizing an argument with an overt (lexical) form. Allen (2000) performed detailed statistical analyses and reported that overt arguments were more likely to carry an informative rather than an uninformative feature. In addition, logistic regression analyses revealed that informativeness features significantly predicted which arguments appeared as overt

versus which appeared as null forms. In a related study, Campbell, Brooks, and Tomasello (2000) analysed pronoun use in English-speaking children aged 2;6 and 3;6 and reported that the children's choice of pronoun (over a null or lexical form) was influenced by the referent's previous mention, indicating sensitivity to the discourse context.

Not only do children show sensitivity to the informativeness features of discourse referents, but also to the principles underlying the distribution of arguments in syntax. Allen and Schröder (2003) and Clancy (1993; 2003) analysed the syntactic distribution of sentence arguments of the same four children acquiring Inuktitut and the same two children acquiring Korean (e.g., Allen, 1997; 2000; Clancy, 1993; 1997). Overall, results revealed that both Inuktitut and Korean children's speech conformed to the four PAS constraints as outlined by Du Bois (1987). Children showed a tendency to produce only one new argument (One New Argument Constraint) and only one lexical argument (One Lexical Argument Constraint) per clause. New arguments (Given A Constraint) and lexical arguments (Non-lexical A Constraint) tended to be located in subject role of intransitive or object role of transitive verbs, but very infrequently in subject role of transitive verbs. Thus, like adults, children too are sensitive to the underlying cognitive and discourse-pragmatic constraints which motivate argument choice and surface syntactic patterning in discourse.

Evaluation of Discourse-Pragmatic Approaches

The studies reviewed suggest that there is a fundamental relationship between an argument's morphological form, its referential status, and its location in surface syntax;

indicating that argument omission and production is not random, but rather highly systematic. By analysing argument choice via a discourse-pragmatic perspective, this three-dimensional relationship becomes apparent.

Because a discourse-pragmatic account does not posit an all-or-nothing type of rule or principle, it does not fall into the trap of postulating a default pattern of either absolute omission or absolute usage. Thus, the inconsistency of omission is adequately explained. A discourse-pragmatic account can also explain the null argument phenomenon, as well as the null subject/null object asymmetry. Arguments that are overt tend to carry more informativeness features than arguments that are omitted and arguments that carry less informativeness features tend to be omitted. Secondly, discourse-pragmatic accounts assume that, unlike the transitive object role, the subject role of transitive verbs is reserved for given and non-lexical, thus less informative, arguments. As such, informative (typically lexical) arguments tend to appear more frequently in object position than in subject position. Thus, the subject position of transitive verbs is the most likely to show argument omission.

Whether children are learning a null argument or an overt argument language, argument omission and production is predicted to be motivated by the same discourse-pragmatic strategies. For instance, as English-speaking children begin to learn that omitted arguments are not a grammatical option in English, pronominal arguments become more frequent, as pronominals function to represent given information in the same way that null arguments do for null argument languages. Thus, the children continue to use the same discourse-pragmatic principles, except in accordance with the grammatical constraints of the target language. More importantly, discourse-pragmatic

accounts postulate that the same mechanism underlies all of argument representation. As such, there is no need for previous grammars to somehow be "unlearned." A theory that does not require unlearning can explain omission and production of arguments in both child and adult grammars. Some data exist at present to suggest that children are sensitive to informativeness features of discourse from as early as the one-word stage, (e.g., Baker & Greenfield, 1988; Greenfield & Smith, 1976; Greenfield & Zukow, 1978), implying the continuity of (discourse-pragmatic) strategies throughout development.

It has been suggested that linguistic structures that are observed crosslinguistically and that appear early in development might be cognitively privileged (Goldberg, 1998). Indeed, the given-new dimension which corresponds to the certainty-uncertainty distinction is a cognitive process that is common to child language learners and adult speakers alike (Greenfield, 1979). Tomasello (2000a) and others have proposed that the early learning of argument structure might derive from very general cognitive abilities that are manifest in many domains of human activity (also Du Bois, 1987; Goldberg, 1998; Tomasello, 1992b). For instance, a scene in which a human protagonist performs an action to effect an outcome on an inanimate patient or recipient is typically paired with transitive syntactic structure; suggesting a cognitive predisposition to encode basic perceptual experiences into basic language patterns. In fact, basic perceptual situations such as these are common experiences for young children crosslinguistically (Tomasello, 1992b; 2000a). As suggested by Chafe (1987; 1994; 1996) and Du Bois (1987), not only are the discourse-pragmatic strategies underlying referential choices cognitive in nature, but there appears to be an upper limit on the amount of information that can be transmitted and decoded at any one time. Thus, a discourse-pragmatic account inherently

assumes a "built-in" processing limitation. Performance limitations on speech production are most likely involved, whether the speaker is a child or adult, given the nature of working memory and its limited resources.

Finally, a discourse-pragmatic approach speaks to the idea that the learning of language cannot succeed without the social and pragmatic context in which the learning takes place. Tomasello (1992b; 2000a) talks about "social cognition," the idea that language learning would not arise without cognitive processes such as joint attention and mutual knowledge (i.e., awareness of the knowledge states of others). Such processes comprise some of the reasons for choosing a particular expression in a particular discourse situation. For instance, in order to choose between a null, pronominal, or lexical argument, the child must assess the specific knowledge states of a specific listener on a specific discourse occasion (Campbell, Brooks, & Tomasello, 2000; also Chafe, 1987; 1994; 1996; Givón, 1983; Skarabela & Allen, 2002); assessment abilities which require skills of social cognition on the part of the (child) speaker. A discourse-pragmatic approach basically takes advantage of the child's developing cognitive and social-cognitive skills – the same cognitive and learning mechanisms which are assumed to operate across development (Tomasello, 1992b; 2000a).

In sum, a discourse-pragmatic perspective provides a theory which takes into account not only the mechanisms for learning discourse-sensitive rules of argument omission and production, but also takes account of the cognitive, memory, and processing constraints that come into play during natural language discourse. These operate together with pragmatic features of discourse to ultimately shape the surface form of utterances.

Justification and Overview of the Thesis

In order to fully support a discourse-pragmatic theory of child argument representation, some crucial data are needed. The only data available at present in support of a discourse-pragmatic explanation of children's argument representation are restricted to studies of null argument languages where argument omission is a grammatical option (e.g., Inuktitut, Korean). It is unclear at this point whether discourse-pragmatic strategies could explain argument choice in children learning overt argument languages or even whether these children show patterns similar to those observed with children learning null argument languages.

Also lacking are developmental studies. The lack of studies investigating the development of discourse-pragmatic strategies makes it difficult to evaluate whether a discourse-pragmatic account can explain the gradual shift towards consistent overt argument usage as is typically observed with children learning overt argument languages, nor whether children learning null argument languages continue to omit arguments with the same motivation across developmental periods or if omitted arguments in later (adult) language are governed by different processes. There has been some attempt to analyse longitudinal null argument data into developmental stages (e.g., Allen & Schröder, 2003), but lacking are data from earlier stages of development (Allen and Schröder's data began at MLU 3.25).

Finally, a systematic investigation of the input children receive is also required; which is at present also lacking. In a social-pragmatic point of view, language is a social act and children learn social acts from imitating parental input (what Tomasello (2000a) refers to as "cultural learning"). The role of input cannot be ignored, as only by analysing

input can we investigate how children learn language-specific patterns of argument representation. Although previous research has observed the use of discourse-pragmatic strategies in adult speech to children (e.g., Clancy, 1993; Fisher & Tokura, 1995), few studies have investigated the relationship of parental input to the development of children's discourse-pragmatic strategies. In addition, whether discourse-pragmatic input patterns are similar across languages of diverse typologies remains an open question at this point in time. Lacking is a crosslinguistic analysis in order to distinguish between universal and language-specific input patterns.

In this thesis, I investigate whether the discourse-pragmatic approach can explain the development of argument representation in child language. I report on three studies, all of which analyse argument choice crosslinguistically and developmentally, as well as the relationship between parental input strategies and children's referential choices. In the first study, the relationship between the referential form (null, pronominal, or lexical) and pragmatic information (given or new) of sentence arguments was analysed in a study of six English-speaking and six Japanese-speaking children when the children were 21 and 36 months of age. In the second study, not only argument form and referential status, but also whether the referring expression was supplemented with non-linguistic pragmatic correlates (pointing, reaching, moving, making a head motion, or purposeful gaze direction toward a referent) were analysed at each of four linguistic periods between MLU 1.00 and 4.00 in two monolingual English-speaking and two monolingual Japanese-speaking children. Finally, in the third study, the relationship between referential form, pragmatic information, and syntactic role (transitive subject, transitive

object, or intransitive subject) was analysed in the same two groups of 21- and 36-month-old children. The motivation for each study is given at the beginning of each chapter.

Table 1
Developmental Trajectory of Argument Omission in Overt, Null, and Null Subject Languages

Language Typology	Language	Design	Summary of Findings
Overt Argument			
P. Bloom (1993); Hyams & Wexler (1993)	English	Longitudinal; from 1;6 to 2;3	subjects omitted 60% at 1;6, 27% at 1;9, 32% at 1;10, 11% at 1;11, 13% at 2;0, 7% at 2;3
Valian (1991)	English	Cross-sectional; four MLU groups (1.77; 2.49; 3.39; 4.22)	overt subjects used 69%, 89%, 93%, 95%; overt objects used 93%, 93%, 98%, 97%
Wang, Lillo-Martin, Best, & Levitt (1992)	English	Cross-sectional; two MLU groups (3.51; 4.48)	subject omission declined from 25.89% to 8.93%; object omission declined from 8.3% to 1.5%
Weissenborn (1992)	French	Longitudinal; from 2;1 to 2;9	overt subject use increased from 69.7% to 93.7%
Weissenborn (1992)	German	Longitudinal; from 1;10 to 2;8	overt subject use increased from 18.4% to 76.6%
Null Argument			
Allen & Schröder (2003)	Inuktitut	Longitudinal; three MLU groups (3.25–3.99; 4.00–4.74; 4.75–5.49)	combined subject and object omission 88%, 82.5%, 74.9%
Hirakawa (1993)	Japanese	Longitudinal; four age periods (2;1, 2;4, 2;9, 3;0)	subjects omitted 90.5%, 89.4%, 91.4%, 86.4%; objects omitted 47.6%, 40%, 30.9%, 40.8%
Kim (2000)	Korean	Longitudinal; three children from 1;7 to 2;6	overt subjects used 30%, 33%, 43%; overt objects used 54%, 62%, 51%
Nakayama (1994)	Japanese	Longitudinal; four age periods (1;0, 1;6, 2;0, 2;6)	subjects omitted 56.5%, 73.7%, 65.7%, 66%; objects omitted 55.5%, 42.7%, 42.3%, 36.3%
Wang, Lillo-Martin, Best, & Levitt (1992)	Chinese	Cross-sectional; three MLU groups (3.41; 4.41; 5.28)	subjects omitted 55.73%, 45.65%, 38.25%; objects omitted 20.2%, 21.3%, 26%
Null Subject			
Austin, Blume, Parkinson, Núñez des Prado, & Lust (1997b)	Spanish	Cross-sectional; individual MLUs ranging from 1.25 to 4.33	subject omission declined from 100% to 50%
Valian (1991)	Italian	Longitudinal; two age periods (1;6–1;10, 2;0–2;5)	overt subjects used 30% both times

Note. Kim (2000) provides proportion rates for each month of recording, but averages are given here for presentation purposes (calculated by this author). Only children P, C, and J are included.

CHAPTER 2

Transition to Manuscript

Previous research showed that young children are sensitive to discourse-pragmatic principles early on in language development (e.g., Allen & Schröder, 2003; Clancy, 2003). However, this research was conducted with children learning null argument languages (e.g., Korean and Inuktitut). In order to examine whether a discourse-pragmatic approach can account for children's referential choices crosslinguistically, a study analysing discourse-pragmatic strategies in children learning overt argument languages is needed. In the manuscript that follows, a crosslinguistic developmental study was conducted by analysing argument representation in monolingual English-speaking and monolingual Japanese-speaking children at 21 and 36 months of age. Referential choices provided in the children's parental input were also analysed. The input analyses were included in order to investigate the influence of parental referential choices on children's learning of language-specific discourse-pragmatic strategies.

Running Head: ARGUMENT REPRESENTATION IN ENGLISH AND JAPANESE

** MANUSCRIPT IN REVISION FOR PUBLICATION**

The Development of Argument Representation in English and Japanese Child Language: A Discourse-Pragmatic Perspective

A M Sonia Guerriero and Yuriko Oshima-Takane

McGill University

Yoko Kuriyama

International Christian University

This research was supported by a Social Sciences and Humanities Research Council of Canada internal grant to the first author and a Natural Sciences and Engineering Research Council of Canada grant to the second author. We thank Amy Cooper, Yuhko Kayama, and Alyssa Ono for their help in transcribing and coding the Japanese data and Fred Genesee, Rachel Mayberry, and Tom Shultz for helpful comments. We especially thank the parents and children who participated in the study.

Address for correspondence: A M Sonia Guerriero, Department of Psychology, McGill University, 1205 Docteur Penfield Avenue, Montréal, Québec, H3A 1B1, CANADA. Email: sonia.guerriero@mail.mcgill.ca.

Abstract

The development of argument representation was examined in a crosslinguistic study of six English-speaking and six Japanese-speaking children when the children were 21 and 36 months of age. Argument choice was investigated using a discourse-pragmatic approach in which the relationship between the referential form (null, pronominal, or lexical) and pragmatic information (given or new) of sentence arguments was systematically analysed. Under a discourse-pragmatic perspective, information which is newly introduced into discourse tends to be lexicalized, whereas information previously mentioned tends to be non-lexicalized (omitted or pronominalized). Results revealed that neither group of children made referential distinctions between given and new information at 21 months of age, suggesting that they were not sensitive to (universal) discourse-pragmatic strategies underlying argument choice. By 36 months of age the English-speaking children's referential distinctions between given and new information were consistent with a discourse-pragmatic strategy. In addition, their referential patterns resembled those of their mothers, indicating that language-specific discourse-pragmatic referential strategies had been learned. At 36 months of age the Japanese-speaking children continued to make inappropriate referential choices, suggesting a violation of discourse-pragmatic strategies. Closer examination revealed a similar referential pattern in the children's input. However, both mothers and children tended to supplement such referential expressions with additional non-linguistic information in order to clarify the referent. These findings are discussed with respect to the learning processes involved in the development of language-specific argument representation.

The Development of Argument Representation in English and Japanese Child Language: A Discourse-Pragmatic Perspective

Introduction

Children's early sentences are replete with omitted, or "null," arguments. Young children's utterances often result in sentences such as the following:[7]

(1) a. Ø want go get it.

b. Ø helping Mommy.

c. Mummy get Ø.

d. Man taking Ø.

e. Ø put Ø on.

These English sentences are lacking either the subject argument (1a and 1b), object argument (1c and 1d), or both (1e). At some point, children need to learn that such utterances are ungrammatical in English and that overt arguments need to be supplied. Null[8] arguments in child language are not restricted to English. In fact, null arguments are observed crosslinguistically, regardless of whether the children are learning an overt argument language, such as English, German, or French, in which the grammar calls for both subject and object arguments to be overtly realized, or a null subject language, such as Italian or Spanish, in which omission of the subject is a grammatical option, or even a

---

[7] Examples are cited from L. Bloom (1970) and Radford (1990). The Ø symbol represents the location of an omitted argument.

[8] We use the terms "null" and "omitted" interchangeably to indicate when a verbal argument is not overtly produced in speech. We do not make any specific statements as to whether or not the omitted argument is a "zero pronoun." The term "argument representation" is used to describe the morphological form of the verbal argument as it appears in surface syntax, whether it appears as a null, pronominal, or lexical form.

null argument language, such as Japanese, Chinese, or Inuktitut, in which either or both subject and object arguments can be freely omitted. In contrast to children learning an overt argument language, children learning a null subject or null argument language can continue to omit arguments as a grammatical option. The purpose of this study is to investigate this null argument phenomenon. First, we investigate whether children's early null arguments are randomly omitted or whether these can be shown to follow a systematic pattern. Secondly, given the preponderance of null arguments at earlier ages, we aim to understand how it is that children come to learn language-specific principles of argument representation, such as whether argument omission is or is not a grammatical option.

Developmental Patterns of Argument Omission

Many studies of children's null arguments have been undertaken, and a few have looked at this phenomenon developmentally (e.g., Allen & Schröder, 2003; Austin, Blume, Parkinson, Núñez del Prado, & Lust, 1997a; Bloom, 1993; Hirakawa, 1993; Hyams & Wexler, 1993; Nakayama, 1994; Valian, 1991; Weissenborn, 1992). The few that have analysed longitudinal data indicate that argument omissions for children learning overt argument languages gradually decline over time until they largely disappear, at which point overt arguments become predominant (e.g., Bloom, 1993; Valian, 1991; Weissenborn, 1992). There is no abrupt or sudden shift from omission to overt usage. Not only is the decline in argument omission gradual, but the developmental pattern is reported to be highly variable. An unswerving linear decline is not usually observed. For instance, Weissenborn (1992) reported the longitudinal data of two

children, one learning French (aged from 2;1 to 2;9) and another learning German (aged from 1;10 to 2;8). The children's language was sampled every few months. The French-speaking child produced overt subjects 69.7%, 63.5%, 70.3%, 91.2%, 87.1%, 95.7%, and 93.7% of the time over the eight-month period. The German-speaking child's overt subject production was 18.4%, 17.4%, 44.1%, 60.6%, 85.9%, and 76.6% of the time over the ten-month period. A similar finding was reported by P. Bloom (1993) and Hyams and Wexler (1993) who analysed subject omission in two of the English-speaking children studied by Roger Brown (Brown, 1973). The children's subject omissions fluctuated over time until they eventually tapered off.

The same decline in argument omission is not observed in children learning null argument languages. Nakayama (1994) analysed child Japanese argument omission and reported that subject and object omission rates did not change much in three children aged from 1;0 to 2;6. The children showed mean subject omission rates of 56.5% at 1;0, 73.7% at 1;6, 65.7% at 2;0, and 66% at 2;6. Objects were omitted an average of 55.5%, 42.7%, 42.3%, and 36.3% at 1;0, 1;6, 2;0, and 2;6, respectively.[9] A comparison group of seven Japanese-speaking adults showed subject omission an average of 67% of the time and object omission an average of 55% of the time, indicating that the children showed omission rates similar to adult Japanese speakers. Other investigators of child Japanese (e.g., Hirakawa, 1993) have similarly observed the maintenance of argument omission across developmental periods. Allen and Schröder (2003) reported developmental data in a group of four children learning Inuktitut, also a null argument language. The children's

[9] Only two children's data comprise the average reported for 1;0.

data were categorized into three MLU groups according to increasing grammatical complexity, MLU 3.25–3.99, MLU 4.00–4.74, and MLU 4.75–5.49. The three groups showed 88%, 82.5%, and 74.9% argument omission rates (subject and object omission combined). Similar studies with null subject languages, such as Spanish and Italian, have also reported the persistence of omission over time (Austin, Blume, Parkinson, Núñez del Prado, & Lust, 1997a; Valian, 1991). In sum, when compared to children learning overt argument languages, children learning null subject or null argument languages do not show a decline in their use of null arguments. They continue to omit arguments across developmental periods.

Explanations of Children's Argument Omissions

Traditionally, explanations of children's null arguments have taken either competence- or performance-based approaches. Competence-based approaches, also called syntactic or linguistic approaches, propose that children's underlying grammatical competence is different from adults' competence and thus allows argument omission where the adults' grammars would not. Later, children's grammars mature or develop into the adult targets, typically based on evidence found in parental input (e.g., de Haan & Tuijnman, 1988; Hamann & Plunkett, 1998; Hoekstra & Hyams, 1996; Hyams, 1992; 1996; 2001; Hyams & Wexler, 1993; Jaeggli & Hyams, 1987; Lillo-Martin, 1994; O'Grady, Peters, & Masterson, 1989; Radford, 1990; 1996; Sano & Hyams, 1994; Wang, Lillo-Martin, Best, & Levitt, 1992; Weissenborn, 1992). However, competence-based approaches typically fail to account for the gradual shift from predominantly omitted arguments to predominantly overt arguments that is observed in children learning overt

argument languages because these would predict a sudden shift from one phase to the next.

Performance-based approaches (e.g., Bloom, 1990; Bloom, 1993; Gerken, 1991; Pinker, 1984; Valian, 1990; 1991; 1994; Valian & Eisenberg, 1996), on the other hand, propose that children's ungrammatical argument omission occurs at the level of speech production, not underlying competence. Children are assumed to represent correct grammatical competence from a very early age, but omit arguments as a result of immature or limited processing resources. As children's processing capacities mature, argument omission gradually declines until it largely disappears. As such, performance-based accounts improve on this weakness of the competence-based accounts. But a performance limitation, however, fails to explain the language behaviour of all children. If a performance factor is at play with children learning overt argument languages, it is unclear what motivates the same phenomenon for children learning null argument languages. These children's rates of omissions do not decline over time, nor are they expected to. Though performance limitations likely do affect young children – that claim is not disputed – it is possible that such limitations per se are not at the root of children's argument omissions.

A third approach, the discourse-pragmatic approach, is one which has shown success in explaining adult argument omission (e.g., Ariel, 1990; 1996; Chafe, 1987; 1994; 1996; Du Bois, 1987; Givón, 1983; Gundel, 1996; Gundel, Hedberg, & Zacharski, 1993; also, Lambrecht, 1994). This approach is different from the competence- and performance-based approaches in that it integrates grammar with pragmatic principles in the understanding of how referring expressions are represented in syntax and,

furthermore, discourse. Research indicates that adult speakers show sensitivity to discourse-pragmatic factors underlying the differential use of referring expressions in discourse (Chafe, 1994; also, Du Bois, 1987). Essentially, the speaker's assumptions made in regards to the status of information in the mind of the listener are reflected in the speaker's choice of referential forms used to express such information in discourse. Given, previously established referents are more likely to be represented by non-lexical (null or pronominal) arguments, whereas referents newly introduced to the discourse are more likely to be represented by lexical (nominal) arguments, basically in the speaker's efforts to be as explicit as possible for the listener to unambiguously identify the referent. In discourse-pragmatic theory, new information is considered to be more informative and more pragmatically prominent than given information. Similar findings have been reported by others investigating the distribution of referring expressions in discourse with adult speakers of typologically diverse languages, such as English (Chafe, 1987; 1994; Gundel, Hedberg, & Zacharski, 1993), Seneca (a Native American language spoken on reservations in western New York State; Chafe, 1987; 1994), Sakapultek Maya (a Mayan language spoken in highland Guatemala; Du Bois, 1987), Hebrew, German, Brazilian Portuguese, French (as cited in Du Bois, 1987), Japanese (as cited in Du Bois, 1987; Gundel, Hedberg, & Zacharski, 1993), Mandarin Chinese, Spanish, and Russian (Gundel, Hedberg, & Zacharski, 1993), among others. Though some of these studies are preliminary and involve spoken narratives rather than spontaneous discourse, the general tendency regarding the expression of referential forms is clearly manifest, regardless of the type of discourse involved.

Given the success of the discourse-pragmatic approach in explaining the distribution of referring expressions in adult language, a few studies have recently investigated the adaptability of the approach to children's referential choices. Clancy (1993; 1997) conducted a year-long study with two children (aged 1;8 and 1;10 at the start) learning Korean, a language that allows null arguments. She specifically analysed null, pronominal, and lexical arguments and their pragmatic prominence. Pragmatic prominence was operationalized by analysing each argument's referent into four discourse features presumed to influence lexical realization: (1) the referent is queried or is an answer to a question, (2) the referent is in a contrastive context (e.g., it is one of at least two potential referents to which the same predicate could be applied), (3) the referent is introduced into discourse for the first time, and (4) the referent is absent from the context at time of mention. A referent with one or more of these discourse features was considered to be pragmatically prominent. Results revealed that lexical arguments were more likely than omitted arguments to contain one or more discourse features, indicating that the children chose to represent a pragmatically prominent referent with a lexical form. Pronominal forms patterned like both omitted and lexical arguments, reflecting the dual nature of Korean pronouns as attenuated, but overt forms. Clancy also investigated the use of referential forms in the speech of the two children's mothers and found similar results: Lexicalized arguments were the most likely, while omitted arguments the least likely, to refer to pragmatically prominent referents.

Similarly, Allen (1997; 2000) analysed referential choice over a nine month period in four children (aged 2;0, 2;6, 2;10, and 2;6 at the start) acquiring Inuktitut. In addition to the four discourse features explored by Clancy (1993; 1997), (1) query, (2)

absence, (3) contrast, and (4) newness, Allen appended a further four: (5) differentiation in context, (6) differentiation in discourse, (7) inanimacy, and (8) third person. Termed "informativeness" features, these functioned in a binary manner: A positive value indicated informativeness, or more pragmatic prominence; a negative value indicated non-informativeness, or less pragmatic prominence. The assumption was that the identity of a referent characterized by one or more informativeness features with a positive value would be less certain and, thus, the referent more informative. Logistic regression analyses revealed that the eight informativeness features as a set reliably predicted which arguments appeared as overt versus null forms. Further analyses revealed that increasing the informativeness value of a referent increased the likelihood of using an overt morphological form.

These studies (e.g., Allen, 1997; 2000; Clancy, 1993; 1997) indicate that children, too, show sensitivity to the informativeness features of discourse referents and provide some preliminary support for a discourse-pragmatic explanation of child null arguments. Where the performance-based explanation fails with respect to null argument languages, the discourse-pragmatic account succeeds: Both Korean and Inuktitut are null argument languages. However, lacking are discourse-pragmatic investigations with children learning overt argument languages, such as English. As for children learning null argument languages, it is not clear at present whether they continue to omit arguments with the same motivation across developmental periods, or if omitted arguments in later (adult) language, or across developmental periods, are governed by different processes. Allen and Schröder (2003) analysed longitudinal null argument data into developmental stages, but data from earlier periods of development are lacking (Allen & Schröder's

(2003) data began at MLU 3.25). Clancy (1993) has made some observations in regards to developmental changes in children's use of discourse-pragmatic strategies, but because her data were not analysed in terms of developmental periods, these changes are not clearly defined. Finally, few studies have investigated the relationship between parental input and the development of children's discourse-pragmatic strategies. Although previous research has observed the use of discourse-pragmatic strategies in adult speech to children (e.g., Clancy, 1993), no studies have investigated the effects of parental input on the development of children's discourse-pragmatic strategies in order to understand how children learn language-specific argument representation.

The Present Study

To address these issues, we conducted a crosslinguistic investigation of the development of argument representation in two groups of monolingual children, English-speaking and Japanese-speaking. We propose that null arguments in early language occur in accordance with pragmatic principles of discourse and that the same discourse-pragmatic strategies are reflected throughout development, including the eventual mastery of language-specific argument representations. We argue that argument omission and production is not random, but follows a systematic pattern and that this pattern is predicated on pragmatic features of discourse referents. The same discourse-pragmatic principles underlie all forms of argument representation, whether the referential choice is an omitted or an overt argument. We further argue that the discourse-pragmatic account can explain argument representation across typologically distinct languages. Our claim is

that these strategies are universal, but that the specific morphological form used is a grammatical property of a particular language.

We examined English and Japanese because Japanese allows optional omission of either subject or object arguments, whereas English requires both to be overt. For example, the Japanese sentences illustrated in 2b, 2c, and 2d below are all grammatical, whereas the English equivalents are all ungrammatical:[10]

(2) a. *Hanako-ga ringo-o kat-ta.*

Hanako-NOM apple-ACC buy-PAST.

b. *Ø ringo-o kat-ta.*

Ø apple-ACC buy-PAST.

c. *Hanako-ga Ø kat-ta.*

Hanako-NOM Ø buy-PAST.

d. *Ø Ø kat-ta.*

Ø Ø buy-PAST.

This grammatical difference between English and Japanese makes them interesting cases for crosslinguistic analyses. Unlike other null argument languages, such as Italian in which recovery of a null subject pronoun is possible due to rich inflectional markings on the verb, Japanese does not mark verbal agreement. Case is marked via the use of case markers. However, an omitted argument cannot be recovered via its case marker because the case marker is also omitted when the argument is omitted. Recovery of the identity of a null argument in Japanese is typically made via discourse and pragmatic contexts

[10] The following abbreviations are used in glosses: Ø = omitted argument; ACC = accusative case marker; NOM = nominative case marker; PAST = past tense marker.

(Wang, Lillo-Martin, Best, & Levitt, 1992). English, in contrast, is a language in which argument omission is not permitted. Case is marked via word order, though some pronouns have nominative and accusative forms. There are but a few overt agreement inflections on the verb, such as the third-person singular *-s* and the past-tense *-ed.*

Based on the discourse-pragmatic account, we hypothesized that the children's referential choices were motivated by pragmatic features of discourse referents. Referential choice was investigated by analysing the relationship between the morphological form (null, pronominal, or lexical) and pragmatic information (given or new) of sentence arguments in the children's and their mothers' language. We expected that given information would be non-lexicalized (either null or pronominal) and new information would be lexicalized, regardless of the speaker's language background. With respect to the null argument phenomenon occurring in early language, we expected these null arguments to represent given information. Analyses were performed at two different time points in the children's linguistic development, at 21 and 36 months of age. We expected that language-specific referential forms for the representation of given information would be learned from parental input, and that this would be evident by 36 months of age. Thus, Japanese-speaking children show omission of given arguments, whereas English-speaking children show pronominalization.

Method

Participants

Six typically-developing English-speaking and six typically-developing Japanese-speaking girls and their mothers participated in the study. All children received either

English-only or Japanese-only input from both parents. All children were first-born and came from middle-class families. The data were originally collected for longitudinal projects on language development. The English-speaking girls and mothers were observed when the children were 21, 24, 30, and 36 months of age (Oshima-Takane, Goodz, & Derevensky, 1996; Oshima-Takane & Oram, 1991) and the Japanese-speaking girls and mothers were observed when the children were 12, 21, 30, 36, and 48 months of age (Kuriyama, 1999). The results reported here are based on data collected when the children were 21 and 36 months of age. At the 21-month observational sessions, the English-speaking children ranged in age from 1;08.29 to 1;10.14; and at the 36-month observations, they ranged from 2;11.30 to 3;01.05. The Japanese-speaking children ranged in age from 1;08.23 to 1;10.17 at the 21-month observational sessions and from 2;11.29 to 3;00.24 at the 36-month observations.

Data Collection

The children and their mothers visited a university playroom that served as an observational room. English-speaking participants were video-recorded in Canada; Japanese-speaking participants in Japan. Each mother-child dyad was individually video-recorded in free-play, naturalistic interaction. Mothers were instructed to play and interact naturally with their children, as they would do at home. The playrooms were furnished with a table and chairs and equipped with various age-appropriate books and toys, such as puzzles, a set of Megablocks, a farm set, dolls, stuffed bears, Mr. and Mrs. Potato Head, Sesame Street puppets, such as Ernie, Bert, and Kermit, a rice-box, a Fisher Price kitchen set, which contained play food such as hamburgers, hotdogs, mustard,

ketchup, assorted vegetables, steak, chicken, eggs, and so forth, and play kitchenware, such as a tea set, pots and pans, plates, drinking glasses, and utensils. The observation rooms were equipped with videocameras, television monitors, and microphones suspended from the ceiling. The persons controlling the recording equipment were instructed to not interact with the children and mothers during recordings.

Play sessions with the English-speaking dyads were recorded for approximately 50 minutes when the children were 21 months of age (two 25-minute sessions with a short break in between) and approximately 25 minutes when they were 36 months of age. For the Japanese-speaking dyads, the play sessions were recorded for approximately 15 minutes when the children were 21 months of age and approximately 25 minutes when they were 36 months of age. For the purpose of this study, the first 15 minutes of each session were used for analysis.

Transcription and Coding

Twelve play sessions (6 English, 6 Japanese) at each of the two age periods (21 months, 36 months) resulted in a total of 24 sessions for analysis. All speech produced by the children and their mothers during the entire recorded play sessions was transcribed according to the CHAT (English) and JCHAT (Japanese) transcription systems (MacWhinney, 2000a; Oshima-Takane, MacWhinney, Sirai, Miyata, & Naka, 1998). Trained research assistants who were native speakers of English or of Japanese transcribed the sessions from the videotapes. To ensure reliability of the transcriptions, all transcripts were verified for accuracy with the videotapes by a second set of trained research assistants. Any disagreements regarding a transcription between the original and

second transcriber were verified by the original transcriber or a third trained research assistant. If an agreement could not be reached, the CHAT/JCHAT convention was used to mark the utterance as a doubtful transcription.

Children's and mothers' language in the first 15 continuous minutes of each of the 24 transcripts was coded and analysed. The English transcripts were coded by the first author. Trained native speakers of Japanese coded the Japanese sample. Coding guidelines and criteria are given in Guerriero, Oshima-Takane, Cooper, Kayama, and Ono (2001).

The unit of analysis is the verbal clause. All transitive and intransitive verbs were analysed. Sentence types included declaratives, exclamatives, imperatives (the null subject of an imperative was not coded), tag questions (the argument appearing in the tag was not coded), and *yes-no* interrogatives. Coding was restricted to nominal referring expressions only, whether these appeared as null, pronominal, or lexical arguments. Our decision to code nominal arguments only was based on the difficulty and subjectiveness of operationally defining and quantifying activities and events, in particular for the coding of referential status (see Du Bois, 1987 for a similar argument). All subject and direct object arguments fitting the above criteria were coded for (1) morphological form and (2) pragmatic information/referential status (i.e., mention). First, each argument was coded for whether it was represented as a (i) null form (e.g., Ø pet the kitty?), (ii) pronominal form (e.g., you pet the kitty?), or (iii) lexical form (e.g., the girl pets the kitty). Second, each argument was coded for whether the referent it represented was (i) "given," if the referent was previously mentioned in the preceding 20 clauses or (ii) "new," if the referent was introduced into discourse for the first time or if it was

mentioned more than 20 clauses previously. First- and second-person pronouns were always coded as "given." Referents present in the situational context, but never referred to, were always coded as "new" upon first mention. Criteria for the coding of referential status followed Chafe (1987) and Du Bois (1987).

The following exclusionary criteria based on guidelines given in Guerriero et al. (2001) were applied: Only arguments appearing within main or matrix clauses were coded. Arguments appearing in relative or embedded clauses, subordinate clauses, *to*-infinitive clauses, or nominal *-ing* participle clauses (e.g., replying "*Sleeping*" in answer to the question "*What is the kitty doing now?*") were not coded, though arguments appearing within two compound clauses were. Arguments referring to activities or events (e.g., I went fishing), including deverbal nouns (e.g., give a kiss to Kermie), were excluded from analysis because they do not refer to concrete referents. All arguments of copular verbs (e.g., the dolly is pretty), complex transitive sentences (e.g., I saw her leave the room), ditransitive verbs (e.g., Mommy put the bear in the box), or of *wh*-questions (e.g., who ate the cookie?) were not considered for analysis. Because analyses were restricted to intransitive and mono-transitive verbs only, arguments of complex transitive and ditransitive verbs were not coded. Arguments appearing within *wh*-questions were not coded because the referent of the *wh*-word could not be determined. Copular verbs were excluded from analysis because these verbs are often omitted in Japanese, which is also a grammatical option. This decision came about because (1) the identification and coding of arguments of omitted verbs becomes subjective; and (2) to maximize possibilities for comparison between the English and Japanese data. Null arguments were differentiated between those that were grammatically acceptable and those that were not,

and grammatical null arguments were not included for analysis. As such, all subject arguments of verbs appearing in the imperative form (e.g., Ø bring the ball here!) were excluded because omitting the subject of an imperative verb is not optional; rather, it is obligatory and grammatically acceptable. Finally, only arguments appearing in spontaneous speech were coded; all those appearing in memorized or routine phrases, constituting unanalyzable chunks, such as social routines (e.g., *excuse me*, *thank you*), songs, poems, and nursery rhymes, were excluded.

To determine the reliability of the coding scheme, two out of six transcripts from each of the English and Japanese samples (33.33%) were chosen at random and independently coded by a second set of trained research assistants, also native speakers of English or Japanese. The reliability coders followed the same guidelines as the original coders and were instructed to not consult and/or discuss the coding with each other. Proportions of agreement were obtained separately for children and mothers and for English and Japanese. Mean percentages of agreement between the original coders and the reliability coders were 86% for the English-speaking children, 87% for the English-speaking mothers, 75% for the Japanese-speaking children, and 79% for the Japanese-speaking mothers.

Measures and Data Analysis

The CLAN programs (MacWhinney, 2000a) were employed to obtain general language measures for each child, such as mean length of utterance (MLU) and total number of intelligible utterances. These measures were based on the first 15 minutes of continuous speech in each transcript. For the data analysis, CLAN programs were used to

obtain frequencies of all combinations of argument form (null, pronominal, lexical) and referential status (given, new) for each child and each mother in each language group, resulting in a total of six frequency measures for each individual speaker (i.e., given null, given pronominal, given lexical, new null, new pronominal, and new lexical arguments). Two additional frequency measures were obtained by combining the frequencies of null and pronominal arguments into a "non-lexical" category (i.e., given non-lexical and new non-lexical arguments).

Results

General Language Measures

The six English-speaking children produced an average of 106.83 intelligible utterances at 21 months of age (ranging from 45 to 144) and an average of 116.83 intelligible utterances at 36 months of age (ranging from 54 to 197). Their mean MLUs were 1.44 (ranging from 1.09 to 2.25) and 3.62 (ranging from 2.90 to 4.15) at 21 and 36 months of age, respectively. The Japanese-speaking children were not very talkative at 21 months and only produced an average 22.33 intelligible utterances (ranging from 8 to 57). However, by 36 months of age, they produced an average of 92.50 intelligible utterances (ranging from 42 to 120). Their mean MLUs were 1.46 (ranging from 1.07 to 1.91) and 3.61 (ranging from 2.86 to 4.29) at 21 and 36 months of age, respectively. Because only the first 15 minutes of speech was used, some children's MLUs are based on less than 100 utterances. Two-tailed *t*-tests revealed a significant difference in total number of intelligible utterances between the English-speaking and Japanese-speaking

children at 21 months of age ($t(6) = 4.63, p < 0.05$), but not at 36 months of age ($t(10) = 0.98, p > 0.05$).[11]

Argument Omission and Production

For the remaining analyses, we compared the frequencies of argument omission and production for each language group, rather than directly comparing the frequencies between the groups. This is because the purpose of the present study was to investigate whether the patterns of argument omission and production are consistent with universal as well as language-specific discourse-pragmatic strategies, rather than investigating crosslinguistic differences in the sample frequencies. It should also be noted that the significant group difference in total number of intelligible utterances at 21 months of age would make between-group comparisons of frequency data less meaningful.

Group Patterns. Mean proportions of the children's use of null, pronominal, and lexical arguments in reference towards given and new information at 21 and 36 months of age are shown in Figure 1. To test our first hypothesis that children's early null arguments represent given information, we performed a one-tailed *t*-test with paired observations comparing the proportion of null arguments used in reference to given information versus the proportion of null arguments used in reference to new information. *T*-tests were run separately for the English-speaking and Japanese-speaking groups at each age period. The comparison was not significant for the English-speaking

[11] Adjusted degrees of freedom were used as variances were found to be heterogeneous.

children at 21 months of age ($t(5) = 0.91, p > 0.05$), indicating that the children did not differentiate the use of null arguments in reference to given versus new information. However, at 36 months of age, English-speaking children omitted arguments significantly more often in reference to given information than new information ($t(5) = 3.66, p < 0.05$). A single-mean *t*-test revealed that the proportion of null arguments at 36 months of age was not significantly different from zero when used in reference to new information ($t(5) = 1.00, p > 0.05$), although still significant when used in reference to given information ($t(5) = 6.65, p < 0.05$).

FIGURE 1 ABOUT HERE

A one-tailed *t*-test with paired observations showed a significant difference in the proportion of null arguments used in reference to given information versus the proportion of null arguments used in reference to new information for the Japanese-speaking children at 21 months of age ($t(5) = 2.99, p < 0.05$). However, because the children made only one reference to new information at this age period, this result does not necessarily demonstrate that they have understood the referential distinction between given and new information. At 36 months, Japanese-speaking children omitted arguments significantly more often in reference to given information than in reference to new information ($t(5) = 3.52, p < 0.05$), indicating differential use of null forms for given versus new information.

Mean frequencies and standard deviations for the children's and mothers' use of non-lexical (null and pronominal arguments combined) and lexical arguments in reference to given and new information are shown in Figure 2 (English speakers) and

Figure 3 (Japanese speakers). To test our second hypothesis that argument form is motivated by the referential status of discourse referents across developmental periods, a 2 (given, new) x 2 (non-lexical, lexical) x 2 (21 months, 36 months) repeated-measures ANOVA was preformed on the children's and mothers' frequency data for each language group.[12]

FIGURES 2 AND 3 ABOUT HERE

The three-way interaction among referential status, argument form, and age was found to be marginally significant for the English-speaking children ($F(1, 5) = 4.62, p = 0.08$). The children showed a tendency to increase their usage of non-lexical forms in reference to given information between 21 and 36 months of age, whereas there was no corresponding increase in their usage of lexical forms. In reference to new information, while there were slightly more non-lexical forms than lexical forms at 21 months, this tendency was reversed at 36 months, at which time the children used more lexical forms than non-lexical forms. The two-way interaction between referential status and argument form was significant ($F(1,5) = 74.22, p < 0.05$), indicating that the children used non-lexical and lexical forms differentially for given and new information. The two-way referential status by age ($F(1, 5) = 0.62, p > 0.05$) and argument form by age ($F(1, 5) = 0.01, p > 0.05$) interactions were not significant. The age main effect was found to be

---

[12] All F values were obtained from log transformations to stabilize variance. The original mean values were reported in the text and figures; however, because a monotonic transformation does not change what is originally measured by the dependent variables, conclusions can be made on the original measures (Ferguson & Takane, 1989).

significant ($F(1, 5) = 26.85, p < 0.05$), indicating that the children's use of arguments increased as they got older (i.e., they produced more verbs).

The three-way interaction among referential status, argument form, and age for the English-speaking mothers approached significance ($F(1, 5) = 5.33, p = 0.07$). The two-way interactions between referential status and argument form ($F(1, 5) = 91.43, p < 0.05$), referential status and age ($F(1, 5) = 10.88, p < 0.05$), and argument form and age ($F(1, 5) = 7.56, p < 0.05$) were all found to be significant. English-speaking mothers showed a difference in the types of references made between the two age periods: They talked about given information more often when their children were 21 months than 36 months. Furthermore, consistent with a discourse-pragmatic strategy, mothers used more non-lexical forms than lexical forms in reference to given information and more lexical forms than non-lexical forms in reference to new information.

A non-significant three-way interaction among referential status, argument form, and age was obtained for the Japanese-speaking children ($F(1, 5) = 1.05, p > 0.05$). The two-way interactions between referential status and argument form ($F(1, 5) = 10.08, p < 0.05$), referential status and age ($F(1, 5) = 22.38, p < 0.05$), and argument form and age ($F(1, 5) = 9.17, p < 0.05$) were all found to be significant. The children produced few arguments overall at 21 months of age. By 36 months of age their argument use increased considerably, at which time they tended to use more given than new arguments and more non-lexical than lexical arguments overall. Consistent with our expectation, the Japanese-speaking children tended to use more non-lexical than lexical forms in reference to given information. However, contrary to our expectation, they produced slightly more non-

lexical than lexical forms in reference to new information. Reasons for this unexpected pattern are discussed later on.

The three-way interaction among referential status, argument form, and age ($F(1, 5) = 7.74, p < 0.05$) and the two-way interaction between referential status and argument form ($F(1, 5) = 25.773, p < 0.05$) for the Japanese-speaking mothers were found to be significant. The two-way interactions between referential status and age ($F(1, 5) = 5.62, p > 0.05$) and argument form and age ($F(1, 5) = 1.65, p > 0.05$) were not significant. Consistent with a discourse-pragmatic strategy, Japanese-speaking mothers used more non-lexical than lexical forms in reference to given information, and they used this pattern when their children were both 21 and 36 months of age. In reference to new information, the mothers used slightly more non-lexical than lexical forms when the children were 21 months of age, however, this pattern was reversed when the children were 36 months of age, at which time the mothers used more lexical than non-lexical forms.

Mean frequencies and standard deviations for the children's and mothers' use of null and pronominal arguments in reference to given information are shown in Figure 4 (English speakers) and Figure 5 (Japanese speakers). To test our third hypothesis that non-lexical (null and pronominal) arguments in reference to given information are used in accordance with language-specific grammatical principles by 36 months of age, we performed a 2 (null, pronominal) x 2 (21 months, 36 months) repeated-measures ANOVA on frequency data for each of the four groups.

FIGURES 4 AND 5 ABOUT HERE

Consistent with our hypothesis, a significant two-way interaction between argument form and age was found for the English-speaking children ($F(1, 5) = 17.54, p < 0.05$). At 21 months of age, null forms were used almost as frequently as pronominal forms in reference to given information, but by 36 months of age pronominal forms were used more frequently than null forms. The two-way interaction between argument form and age for the English-speaking mothers was not significant ($F(1, 5) = 3.74, p > 0.05$). Consistent with an English-specific discourse-pragmatic strategy, the mothers used more pronominal forms than null forms in reference to given information across age periods, and this main effect was significant ($F(1, 5) = 78.17, p < 0.05$).

A significant two-way interaction between argument form and age was obtained for the Japanese-speaking children ($F(1, 5) = 12.48, p < 0.05$). The Japanese-speaking children used only a few null forms in reference to given information and never used pronominal forms at 21 months of age. However, consistent with our hypothesis, they used null forms most of the time at 36 months of age. The two-way interaction between argument form and age was not significant for the Japanese-speaking mothers ($F(1, 5) = 0.12, p > 0.05$), but a significant main effect of argument form ($F(1, 5) = 316.16, p < 0.05$) was obtained. Consistent with a Japanese-specific discourse-pragmatic strategy, Japanese-speaking mothers used more null forms than pronominal forms in reference to given information across age periods.

Individual Patterns. In order to examine individual patterns of argument omission and production within each language group, frequencies for each combination of referential status and argument form are presented by child in Tables 1 through 4. A child is considered to be sensitive to language-specific discourse-pragmatic strategies if she showed the following patterns: (1) Appropriate reference towards new information, regardless of language, requires more frequent use of lexical forms relative to null and pronominal forms. (2a) Appropriate reference towards given information requires more frequent use of pronominal forms relative to null and lexical forms for an English-speaking child. (2b) In contrast, appropriate reference towards given information requires more frequent use of null forms relative to pronominal and lexical forms for a Japanese-speaking child.

As shown in Table 1, the majority of pronominal forms produced by the English-speaking children at 21 months of age were made by one particular child, E3, and her use of argument forms indicates that she was sensitive to an English-specific discourse-pragmatic strategy. This child's MLU also indicates that she was more linguistically advanced than the other children. The remaining five English-speaking children did not show sensitivity to discourse-pragmatic strategies as their references to given information were made with mainly null forms. Their referential patterns changed considerably by 36 months of age. As shown in Table 2, four of the six children showed an English-specific discourse-pragmatic strategy whereby given information was predominantly pronominalized and new information was predominantly lexicalized. The remaining two children, E2 and E3, did not show a clear pattern. E3's pattern is rather surprising, given that she was clearly showing an English-specific pattern at 21 months of age.

TABLES 1 AND 2 ABOUT HERE

As shown in Table 3, Japanese-speaking children rarely used arguments at 21 months of age. The majority were used by one particular child, J4. J4's MLU at 21 months was slightly higher than the other children's and she was the child who produced the largest number of intelligible utterances. However, whether this child showed sensitivity to a Japanese-specific discourse-pragmatic strategy is questionable. She produced only one new argument appropriately in a lexical form, but as such, it is difficult to determine her ability to differentially make use of the available referential options in representing new information. Three other children, J1, J3, and J6, show a similar non-differentiated pattern due to a low frequency of argument use. At 36 months of age only one child, J4, showed clear differentiation between given and new information whereby given information was omitted and new information was lexicalized (see Table 4). Interestingly, J4 did not have the highest MLU. The remaining five children did not predominantly use lexical forms in reference to new information, though given information was appropriately omitted.

TABLES 3 AND 4 ABOUT HERE

Discussion

Contrary to our expectation that early argument omission would reflect given information, our results indicate that the English-speaking children at 21 months of age did not differentiate the use of null forms to given versus new information. Results for the Japanese-speaking children are inconclusive because, although they did show the expected pattern of null forms used in reference to given information, they rarely produced new arguments at 21 months of age and, thus, it is not possible to ascertain whether they understood the referential distinction between given and new information. These findings suggest that our English-speaking and Japanese-speaking children's early null arguments were not systematically omitted and, furthermore, that the children were not sensitive to discourse-pragmatic strategies underlying referential choice. By 36 months of age, however, both groups of children showed the expected discourse-pragmatic patterns: Given referents were omitted significantly more often than new referents. While null arguments are ungrammatical in English, the children used null forms predominantly in reference to given information and rarely in reference to new information. Closer inspection of their mothers' data revealed that the mothers occasionally used null arguments in reference to given information when the children were 21 months of age. This input pattern could have been a contributing factor for the children's persistent use of null arguments at 36 months of age and is presently under investigation.

Our second expectation, that children's argument choices are motivated by the referential status of discourse referents and that such patterns are observed across developmental periods, received partial support. We expected that the children would use

non-lexical forms in reference to given information and lexical forms in reference to new information and that these strategies would be evident from as early as 21 months of age. Contrary to our expectation, however, our results revealed that the English-speaking children at 21 months of age tended to use more non-lexical, rather than lexical, forms in reference to new information, suggesting that they did not know how to refer to new information appropriately. By 36 months, however, the children's referential choices reflected the use of appropriate discourse-pragmatic strategies: More lexical than non-lexical forms were used in reference to new information. The Japanese-speaking children made only one reference to new information at 21 months of age and this one instance was made with a lexical form. At 36 months of age, however, they tended to use slightly more non-lexical forms than lexical forms in reference to new information, similar to the English-speaking children's pattern at 21 months of age. This is contrary to our expectation and suggests that even by 36 months of age our Japanese-speaking children had yet to learn the appropriate discourse-pragmatic referential strategies.

Finally, we hypothesized that the children's use of non-lexical forms in reference to given information would conform to language-specific morphological forms by 36 months of age, and this hypothesis is well supported by the data. Results showed that, unlike their patterns at 21 months of age, the children's patterns at 36 months of age resembled their respective parental input: Given information was pronominalized by the English-speaking children, whereas it was omitted by the Japanese-speaking children. The similarity between the children's patterns and those of their input suggests that the children had learned the language-specific grammatical and pragmatic rules for making proper reference towards given information.

Our findings call into question our assumption of children's sensitivity to discourse-pragmatic strategies in early stages of linguistic development. Our results seem to indicate that English-speaking and Japanese-speaking children are not sensitive to pragmatic principles of discourse when making referential choices early on in development. Whereas the English-speaking children showed sensitivity by 36 months of age, the same cannot be said for our Japanese-speaking children, who continued to make pragmatically inappropriate referential choices even at 36 months of age. They used non-lexical forms more frequently than lexical forms in reference to new information. This result in particular seems to be at odds with those from previous studies, such as Allen (1997; 2000; 2003) and Clancy (1993; 1997), which have shown that children learning null argument languages are sensitive to the informativeness status of discourse referents. Although it is difficult to compare MLU across different languages, the children in Allen's studies could have been more linguistically advanced than the children in our study. The least linguistically advanced group in Allen's studies ranged in MLU from 3.25 to 3.99.[13] The children in Clancy's studies were likely more comparable to the children in our study. In fact, Clancy observed that one of the Korean-speaking children in her study did not distinguish the use of null arguments between given and new referents. According to Clancy, this particular child was still in the early stages of learning the pragmatic correlates underlying the referential options available in Korean. This observation is in keeping with the findings in our study.

[13] These figures represent what Allen calls "verbal MLU," an MLU based only on utterances containing a verb.

The question remains as to why the Japanese-speaking children in our study were not yet using the appropriate referential strategies even by 36 months of age. Their MLU values at this age suggest that they were not a linguistically atypical group (see Table 4), yet they used non-lexical forms in reference to new information. Closer inspection of the children's mothers' data suggests that the children might have learned this referential pattern from the input. The Japanese-speaking mothers showed a tendency to use more non-lexical than lexical forms in reference to new information when the children were 21 months of age. However, these non-lexical forms in reference to new information tended to represent referents that were present in the situational context and tended to be accompanied by a gesture, such as pointing or touching. Examination of the children's data revealed that the children similarly tended to supplement new arguments with gestures when these were omitted in production. In fact, the use of non-lexical forms in representing new information has been observed by other child language researchers. Skarabela and Allen (2002) report a comparable finding to ours with child speakers of Inuktitut. The children showed a similar tendency to omit new arguments, particularly when these arguments were accompanied by joint attention. These studies, as well as our adult data, show that the use of non-lexical forms in representing new information is not such an atypical phenomenon, at least for speakers of null argument languages. Presumably, non-lexical forms in reference to new information are acceptable in null argument languages if such forms are supplemented with non-linguistic information or if the referent is available in the situational context.

Furthermore, and more importantly to our research objectives, this seemingly aberrant finding with our group of Japanese-speaking children suggests that children's

learning of language-specific argument representation is a two-step process: First, they must learn the referential options available to them in their language, such as whether or not their language will permit null arguments, and secondly, they must master the pragmatic principles underlying the appropriate use of the available referential options. Our findings suggest that the Japanese-speaking children's use of non-lexical forms was not atypical, but that they had yet to master the appropriate referential options for new information. The children were still using mostly null forms for new information when lexical forms would have been more appropriate. It seems that the task for the Japanese-speaking children was made doubly difficult in comparison to that of the English-speaking children. Whereas English-speaking children have two referential options to learn (pronominal and lexical forms), Japanese-speaking children have three (null, pronominal, and lexical forms). Furthermore, the children not only have to pay attention to the morphological forms used, but also to the non-linguistic pragmatic correlates of each form, such as gestures, and to the situational context within which the referent is located. This second component of the learning process becomes quite crucial, particularly for Japanese-speaking children, because of the unconventional, though pragmatically acceptable, use of null forms to represent new information. Given that this is an acceptable referential option in Japanese, we speculate that the Japanese-speaking children's learning of the use of lexical forms to represent new information may have been hampered by the availability of two possible referential strategies for the representation of new information: (1) the use of a null form supplemented by non-linguistic (gestural) information or (2) the use of a lexical form, with or without accompanying non-linguistic information. Our data indicate that the Japanese-speaking

children had yet to learn this second referential strategy for the representation of new information.

How discourse-pragmatic principles come to be is still an open question. Several alternatives are possible. Discourse-pragmatic principles could be available to children from very early on, but they are not used. Perhaps such skills do not manifest themselves until certain prior linguistic or cognitive skills are mastered. Another possibility is that children must learn all discourse-pragmatic principles from "scratch." Such an option may or may not assume prior underlying skills. There is a third possibility, namely that discourse-pragmatic principles are available to children from very early on, but that they are in a primitive, or rudimentary, form and need to be fine-tuned over time until they conform to language-specific principles underlying the appropriate use of referential expressions in discourse. This proposal would accord with others' findings, in particular those of Greenfield and Smith (Baker & Greenfield, 1988; Greenfield & Smith, 1976). Greenfield and Smith conducted a developmental study with two English-speaking children in the one-word stage and reported that the children verbalized the most informative element of the discourse situation (new information) and left unexpressed those elements that were taken for granted (given information). The current state of affairs, however, does not allow us to choose among these three learning alternatives, although the importance of parental input in the learning of language-specific argument representation cannot be understated in this linguistic endeavour.

Also unclear at present is when such discourse-pragmatic strategies come to be fully and properly used. Our English data indicate some time before three years of age for English-speaking children, whereas our Japanese data indicate not before three years of

age for Japanese-speaking children. The precise developmental timecourse for discourse-pragmatic strategies is a question for future studies. In addition, given the variability in MLU among the children at 21 and 36 months of age, a more suitable methodology would be to group children by MLU, rather than age. Conducting a developmental study based on MLU rather than age would more accurately describe the onset and eventual mastery of discourse-pragmatic strategies.

Table 1
Individual Frequencies: English-Speaking Children at 21 Months of Age

Child	MLU	Total Intelligible Utterances	Referential status	Argument Form		
				Null	Pronominal	Lexical
E1	1.52	144	Given	10	1	1
			New	2	1	0
E2	1.23	96	Given	6	0	0
			New	1	0	1
E3	2.25	139	Given	8	48	15
			New	1	2	5
E4	1.09	45	Given	1	0	0
			New	3	0	0
E5	1.36	141	Given	17	2	1
			New	4	0	1
E6	1.20	76	Given	13	3	0
			New	5	0	0

Table 2
Individual Frequencies: English-Speaking Children at 36 Months of Age

Child	MLU	Total Intelligible Utterances	Referential status	Argument Form		
				Null	Pronominal	Lexical
E1	3.43	54	Given	1	16	3
			New	0	2	5
E2	2.90	76	Given	1	10	2
			New	0	4	3
E3	4.06	197	Given	11	55	6
			New	0	19	13
E4	3.31	105	Given	3	29	2
			New	1	3	7
E5	3.88	155	Given	8	76	5
			New	0	6	17
E6	4.15	114	Given	6	44	2
			New	0	7	10

Table 3
Individual Frequencies: Japanese-Speaking Children at 21 Months of Age

Child	MLU	Total Intelligible Utterances	Referential status	Argument Form		
				Null	Pronominal	Lexical
J1	1.07	15	Given	1	0	0
			New	0	0	0
J2	1.63	16	Given	0	0	0
			New	0	0	0
J3	1.13	8	Given	1	0	0
			New	0	0	0
J4	1.91	57	Given	7	0	5
			New	0	0	1
J5	1.59	22	Given	0	0	0
			New	0	0	0
J6	1.44	16	Given	3	0	0
			New	0	0	0

Table 4
Individual Frequencies: Japanese-Speaking Children at 36 Months of Age

Child	MLU	Total Intelligible Utterances	Referential status	Argument Form		
				Null	Pronominal	Lexical
J1	4.29	119	Given	38	3	13
			New	2	7	5
J2	3.67	100	Given	20	1	2
			New	3	0	2
J3	3.35	68	Given	44	1	4
			New	5	0	2
J4	3.43	120	Given	26	2	5
			New	1	0	6
J5	2.86	42	Given	13	0	5
			New	3	1	2
J6	4.04	106	Given	24	3	15
			New	5	2	3

Figure Captions

Figure 1. Mean proportions of null, pronominal, and lexical arguments in reference to given and new information for English-speaking and Japanese-speaking children. (Mean proportions for given information for the Japanese-speaking children at 21 months of age are based on $N = 6$. However, of the six children, only four used null forms and only one child used lexical forms. Mean proportions for new information are not shown because only one Japanese-speaking child produced one new argument at 21 months of age. See Table 3.)

Figure 2. Mean frequencies (and standard deviations) of non-lexical and lexical arguments in reference to given and new information for English-speaking children and mothers.

Figure 3. Mean frequencies (and standard deviations) of non-lexical and lexical arguments in reference to given and new information for Japanese-speaking children and mothers.

Figure 4. Mean frequencies (and standard deviations) of null and pronominal arguments in reference to given information for English-speaking children and mothers.

Figure 5. Mean frequencies (and standard deviations) of null and pronominal arguments in reference to given information for Japanese-speaking children and mothers.

Figure 1

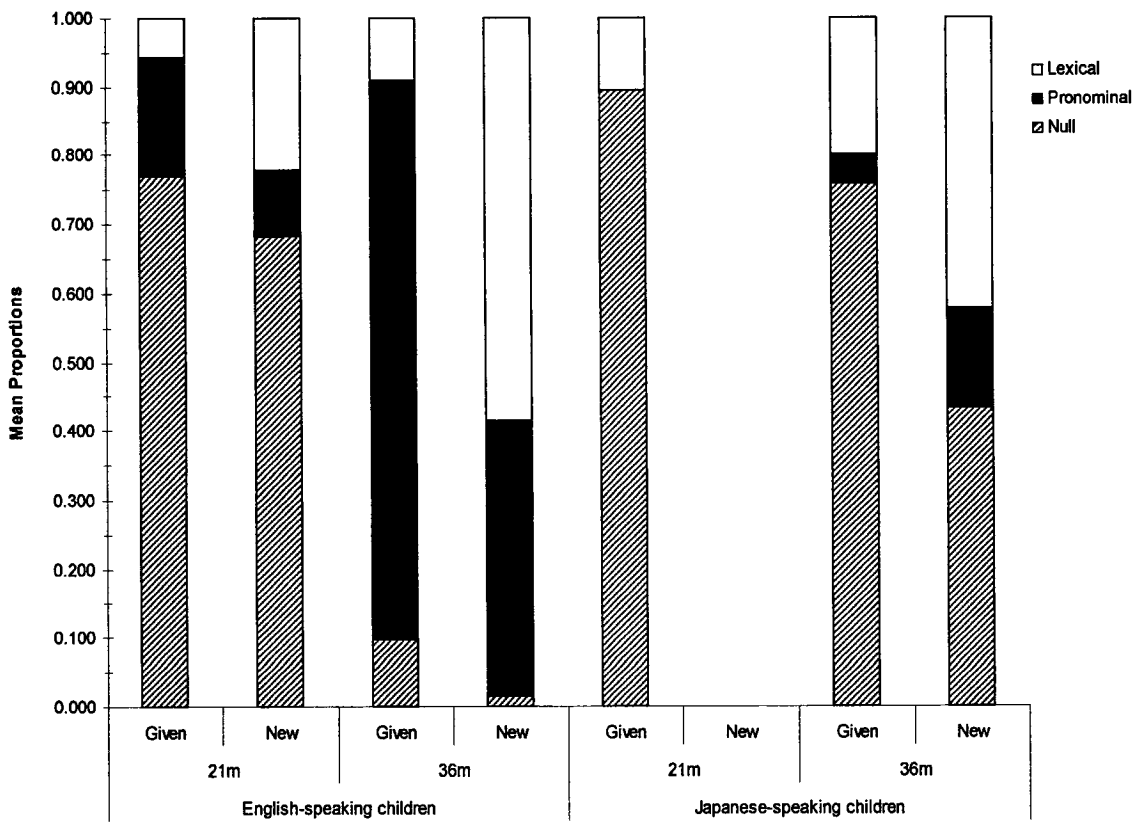

1.000
0.900
0.800
0.700
0.600
0.500
0.400
0.300
0.200
0.100
0.000
Mean Proportions
Lexical
Pronominal
Null
Given
New
Given
New
Given
New
Given
New
21m
36m
21m
36m
English-speaking children
Japanese-speaking children

Figure 2

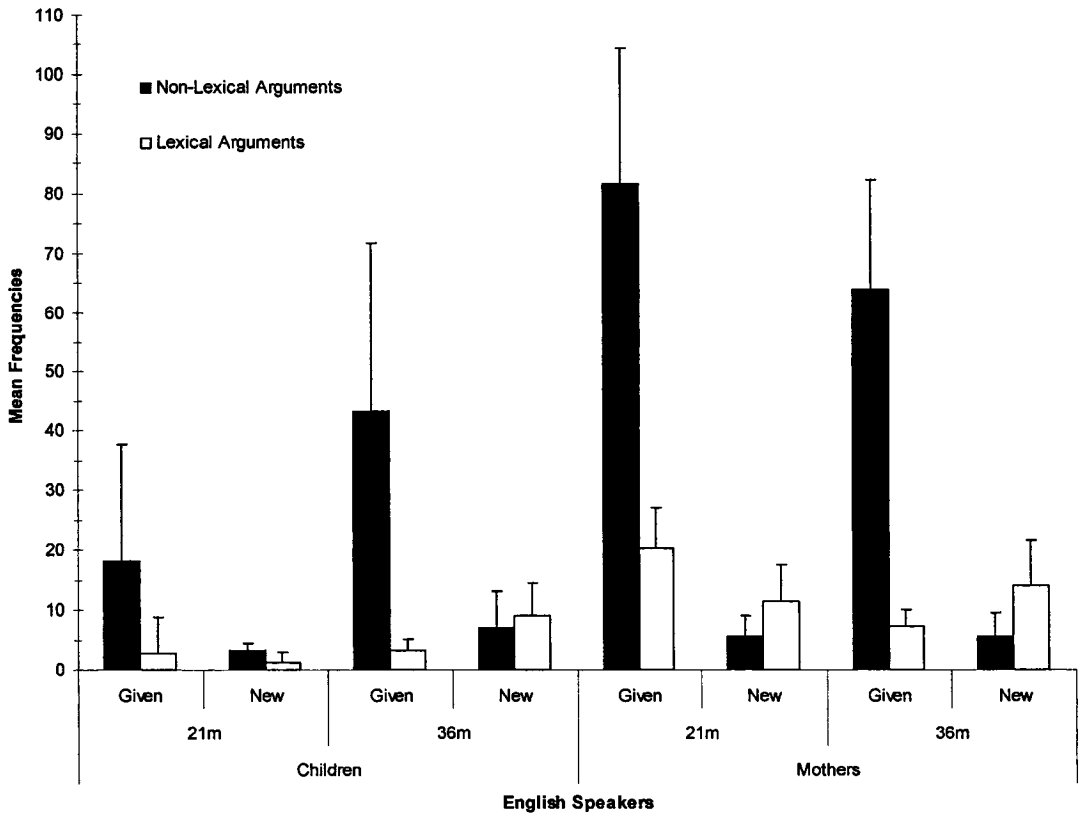

110
100
90
80
70
60
50
40
30
20
10
0
Mean Frequencies
Non-Lexical Arguments
Lexical Arguments
Given
New
Given
New
Given
New
Given
New
21m
36m
21m
36m
Children
Mothers
English Speakers

Figure 3

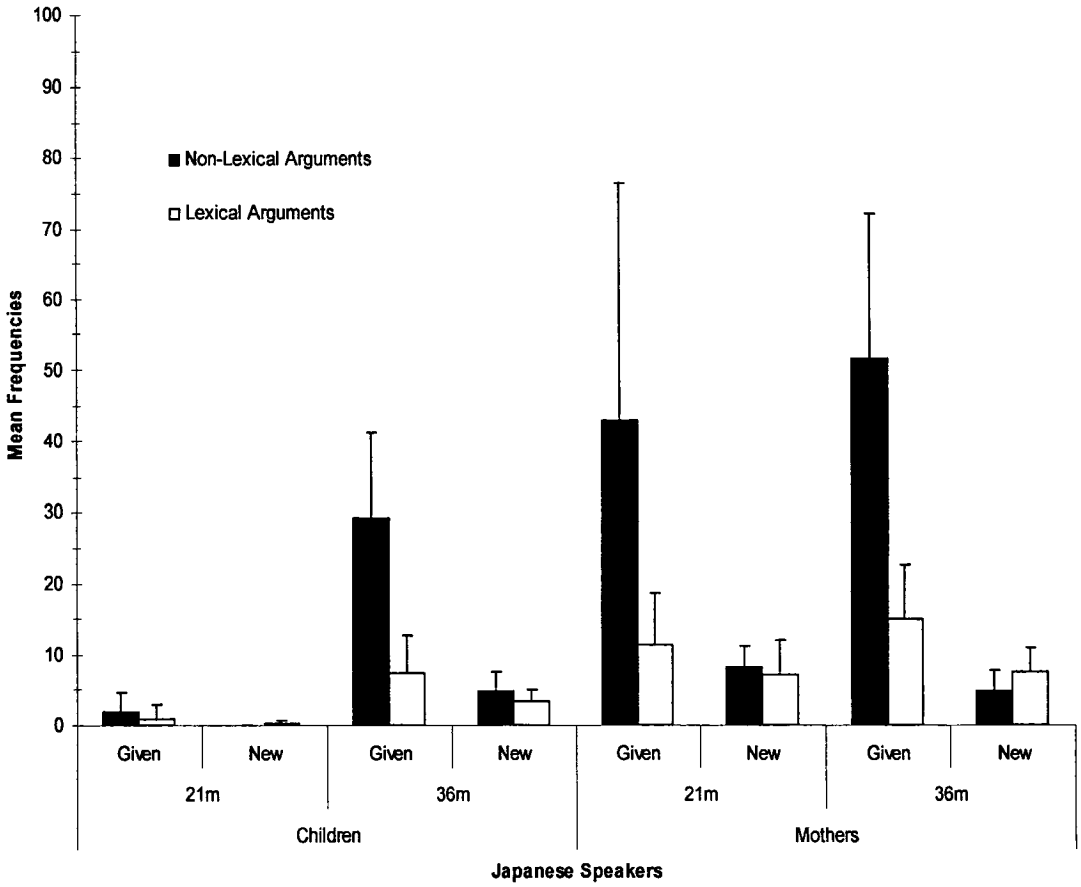

100
90
80
70
60
50
40
30
20
10
0
Mean Frequencies
Non-Lexical Arguments
Lexical Arguments
Given
New
Given
New
Given
New
Given
New
21m
36m
21m
36m
Children
Mothers
Japanese Speakers

Figure 4

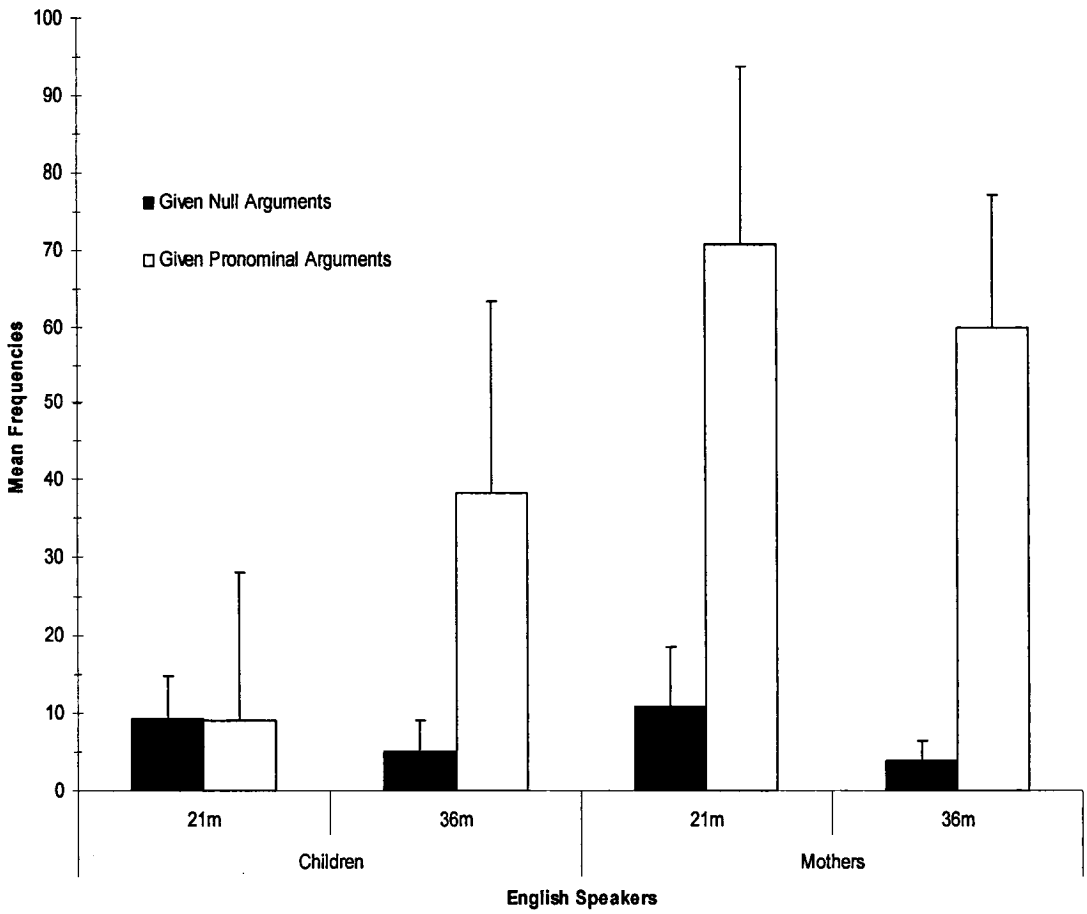
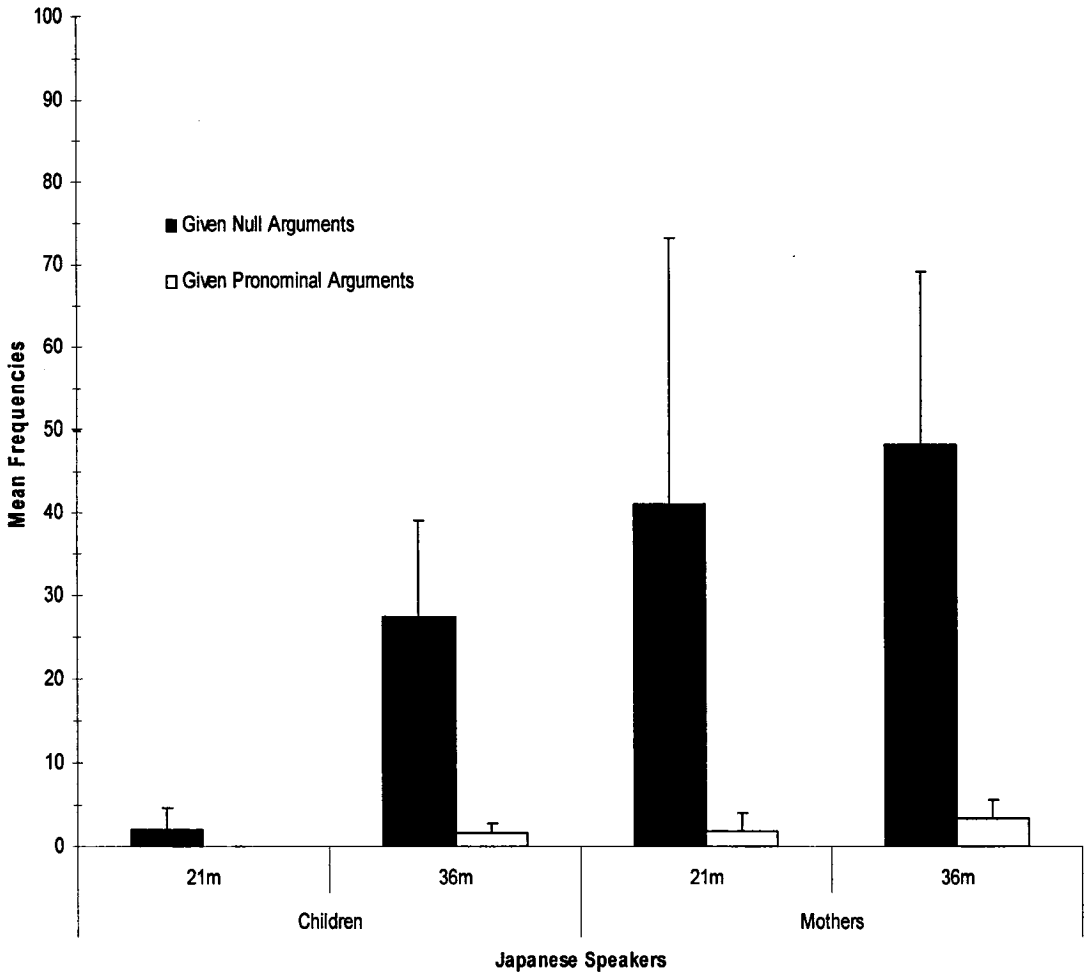

100
90
80
70
60
50
40
30
20
10
0
Mean Frequencies
Given Null Arguments
Given Pronominal Arguments
21m
36m
21m
36m
Children
Mothers
English Speakers

Figure 5


100
90
80
70
60
50
40
30
20
10
0
Mean Frequencies
Given Null Arguments
Given Pronominal Arguments
21m
36m
21m
36m
Children
Mothers
Japanese Speakers

CHAPTER 3

Transition to Manuscript

The previous study revealed that the English-speaking children showed an English-specific referential strategy by 36 months of age, but that the Japanese-speaking children continued to make referential choices inconsistent with expected discourse-pragmatic strategies. The findings also revealed individual variability as a result of grouping children by age. In the study that follows, children are grouped by MLU and referential choices are analysed at four time points in development between MLU 1.00 and 4.00. In addition, the children's use of non-linguistic pragmatic correlates, such as gestures, are also analysed in order to investigate whether children use gestures to supplement informative referents represented with less specific argument forms. Like the previous study, parental input is also analysed in order to investigate the influence of input on children's referential choices, whether linguistic or non-linguistic.

Running Head: CHILDREN'S REFERENTIAL CHOICES

MANUSCRIPT IN PREPARATION FOR SUBMISSION

Children's Referential Choices:

The Role of Non-Linguistic Pragmatic Information

A M Sonia Guerriero and Yuriko Oshima-Takane

McGill University

This research was supported by a Social Sciences and Humanities Research Council of Canada internal grant to the first author and a Natural Sciences and Engineering Research Council of Canada grant to the second author. We thank Amy Cooper, Yuhko Kayama, and Alyssa Ono for their help in transcribing and coding the Japanese data and Fred Genesee, Rachel Mayberry, and Tom Shultz for helpful comments. We are grateful to Susanne Miyata for making the videotapes of the Japanese sessions available to us. We especially thank the parents and children who participated in the study.

Address for correspondence: A M Sonia Guerriero, Department of Psychology, McGill University, 1205 Docteur Penfield Avenue, Montréal, Québec, H3A 1B1, CANADA. Email: sonia.guerriero@mail.mcgill.ca.

Abstract

This study reports on the development of referential choice in two monolingual English-speaking and two monolingual Japanese-speaking children and their mothers. The relationship between the argument form (null, pronominal, or lexical) and referential status (given or new) of sentence arguments, as well as whether the argument was supplemented with non-linguistic pragmatic correlates (e.g., pointing, reaching, gaze direction), was systematically analysed at each of four linguistic periods between MLU 1.00 and 4.00. Results revealed that neither the English-speaking nor the Japanese-speaking children showed sensitivity to discourse-pragmatic principles early on in linguistic development. One English-speaking child began to show an English-specific pattern by Period II (between MLU 2.00 and 3.00); the other by Period III (between MLU 3.00 and 4.00); however, both children showed sensitivity to the referential status of discourse referents from as early as Period I (between MLU 1.00 and 1.99) by using non-linguistic pragmatic correlates to indicate an informative discourse referent. The children's patterns were strikingly similar to the patterns observed in their parental input. The Japanese-speaking children, on the other hand, showed discourse-pragmatic violations as late as Period IV (MLU 4.00 and above), as well as inconsistent use of non-linguistic pragmatic correlates. However, similar (presumed) discourse-pragmatic violations and inconsistent use of non-linguistic pragmatic correlates were also observed in the children's parental input. Furthermore, individual variability in the children's patterns was accounted for by variability in their respective input patterns. Findings are discussed with respect to the importance and influence of input in children's learning of language-specific discourse-pragmatic patterns for argument representation.

Children's Referential Choices:

The Role of Non-Linguistic Pragmatic Information

Introduction

Children at early stages of language development tend to omit sentence arguments from overt production. These early omissions can occur in subject or object position and persist until about 3;0 or 4;0 years of age (Allen, 1997; Guerriero, Oshima-Takane, & Kuriyama, 2004). This "null argument" phenomenon occurs crosslinguistically, whether children are learning an overt or a null argument language. That is, both children who are learning an overt argument language such as English, in which sentence arguments are obligatorily overt, or a null argument language such as Japanese, in which sentence arguments can be optionally omitted, show the same phenomena. One difference exists, however, in that children learning overt argument languages eventually begin to supply overt sentence arguments more and more consistently, whereas children learning null argument languages continue to omit them (Allen & Schröder, 2003; Austin, Blume, Parkinson, Núñez del Prado, & Lust, 1997b; Bloom, 1993; Guerriero, Oshima-Takane, & Kuriyama, 2004; Hirakawa, 1993; Hyams & Wexler, 1993; Nakayama, 1994; Valian, 1991; Weissenborn, 1992). Our overall goal in this paper is to understand how, given the preponderance of omitted arguments in early language, children come to learn language-specific properties of argument representation.

Using a discourse-pragmatic approach (see Chafe, 1987; 1994; 1996; Du Bois, 1987; Givón, 1983), some researchers have shown that children's argument omissions, even at such early stages of language development, follow a systematic pattern. Clancy

(1993; 1997) conducted a year-long study with two children (aged 1;8 and 1;10 at the start) learning Korean, a language that allows null arguments. Results showed that the children were more likely to omit an argument if it represented a referent of lesser pragmatic prominence. By contrast, an argument representing a referent that was pragmatically prominent was more likely to be overtly produced. Pragmatically prominent referents included those that were queried or an answer to a question, in a contrastive context, introduced into discourse for the first time, or absent from the context at time of mention – discourse features presumed to influence overt mention of a referent. Allen (1997; 2000) continued along this line of research and analysed referential choice over a nine-month period with four children (aged 2;0, 2;6, 2;10, and 2;6 at the start) acquiring Inuktitut, also a null argument language. Logistic regression analyses revealed that a set of eight discourse features (newness, contrast, inanimacy, query, absence, differentiation in context, differentiation in discourse, and third person) reliably predicted which arguments appeared as overt versus which appeared as null forms. Children tended to omit less informative arguments (those representing referents characterized by fewer discourse features), but realized as overt more informative ones (those representing referents characterized by more discourse features). Clancy's and Allen's studies were the first to show that young children learning null argument languages are sensitive to pragmatic features of discourse and, moreover, that their referential expressions are not merely random choices.

More recently, Guerriero, Oshima-Takane, and Kuriyama (2004) showed that a discourse-pragmatic account can explain argument choice in an overt argument language as well: English-speaking children aged 36 months (3;0) similarly tended to lexicalize

new referents and non-lexicalize (pronominalize) given ones. However, this study, in which data was analysed developmentally, also revealed that discourse-pragmatic principles might not be fully available or used by children at early periods of linguistic development. The children tended to use more non-lexical (null or pronominal) forms, rather than lexical forms, in reference to new information at 21 months of age (1;9), contrary to the predictions of the discourse-pragmatic account. The results furthermore indicated that the referential choices made by the Japanese-speaking group of children seemed to contradict established discourse-pragmatic expectations (e.g., Allen, 1997; 2000; Clancy, 1993; 1997), even at 36 months of age. The children tended to use slightly more non-lexical (null) than lexical forms to refer to new information. Interestingly, a similar pattern was also observed with their Japanese-speaking mothers. The mothers tended to use more non-lexical than lexical forms in reference to new information when the children were 21 months of age, a pattern which their children reproduced at 36 months of age. Further analyses revealed that the Japanese-speaking children and mothers tended to use gestures, such as pointing to or touching the referent, to clarify such null forms used in reference to new information. Based on these findings, the authors suggested that Japanese-speaking children might after all be sensitive to pragmatic principles of discourse from a very early age, even though the referential expressions chosen might seem inappropriate.

The use of gestures might thus be an important factor in the study of children's referential choices, and previous research has indicated that young children are capable of using gestures in referential communicative situations. For instance, Wilcox and Howse (1982) investigated English-speaking children's use of gestures in communicative

misunderstandings and reported that the children (ranging in age from 1;5 to 2;0) tended to persist in their use of gestures when their requests were misunderstood, even though they were capable of giving a verbal clarification. Pechmann and Deutsch (1982) similarly investigated the use of gestures with Dutch-speaking children (ranging in age from 2;0 to 9;0). Their study showed that the group of two-year-old (2;0) children used a pointing gesture to indicate a preferred object 70% of the time (although their pointing was deemed to be ineffective because the objects from which one was to be selected were intentionally placed extremely close to each other). Using a similar paradigm, a recent study by O'Neill and Topolovec (2001) demonstrated that English-speaking children (ranging in age from 2;6 to 2;10) showed sensitivity to the referential inefficiency of their pointing gestures and were able to adapt their communication accordingly by supplying a nominal descriptor label. Lastly, and of particular relevance to the current study, an early study by Tomasello, Anselmi, and Farrar (1985) with English-speaking children (ranging in age from 1;8 to 3;8) showed that the children tended to use gestures more often when using a pronominal form than when using a lexical form in response to an adult's query for clarification, indicating that they were sensitive to the lower specificity and informational value of pronominal arguments.

The studies cited above provide evidence that children are sensitive to the referential nature of their gestures. It is not unexpected, then, that they may make use of gestures as a pragmatic strategy when making referential choices. It has been shown that 12-month-old infants understand the intention of referential pointing (Thoermer & Sodian, 2001) and that children between 16 and 24 months use gestures (in particular, pointing) as a communicative tool (Guidetti, 2002). Thus, children might use non-

linguistic information in situations in which they do not always conform to discourse-pragmatic principles, such as when they use null arguments in reference to new information. By definition, references to new information are first-time, never-before mentions, necessitating an unambiguous referential expression, such as a lexical form (Chafe, 1994; Du Bois, 1987). If a lexical mention is not used, then an alternate referential strategy would need to be employed, such as the use of a non-linguistic pragmatic strategy, in order for the speaker to adequately point out the new referent to the listener. Likewise, young children might resort to using gestures whenever the appropriate referential forms elude them. The idea that children resort to using gestures when referential choices contradict discourse-pragmatic expectations was alluded to by Guerriero, Oshima-Takane, and Kuriyama (2004), but because a controlled analysis of the relation between choice of argument form and gesture use, if any, was not conducted, this idea remains untested at this time.

A second disconcerting issue that comes up in Guerriero, Oshima-Takane, and Kuriyama's (2004) study is that they report a large degree of variability among individual children. Because children were grouped by age, individual MLUs varied considerably within each age group, and such variability could have contributed to the inconsistent and unexpected findings. In addition, some children hardly produced any subject or object arguments, while others produced a great deal. For instance, only one English-speaking child at 21 months of age showed evidence of a discourse-pragmatic strategy; the remaining five did not. Incidentally, this was the child with the highest MLU at that age. At 36 months, four of the six children showed an English-specific discourse-pragmatic pattern, whereby given arguments were pronominalized and new arguments were

lexicalized. In regards to the group of Japanese-speaking children, only one child produced a fair number of arguments at 21 months of age, but the results were inconclusive because she produced only one new argument. At 36 months of age, this same child was the only one out of the six to show a Japanese-specific discourse-pragmatic strategy, whereby given arguments were omitted and new arguments were pronominalized, though she was not the child with the highest MLU at that age. Furthermore, although Guerriero, Oshima-Takane, and Kuriyama's study was developmental in nature, data were lacking between 21 and 36 months of age, and as such the precise point in time at which children master discourse-pragmatic skills is still unclear.

The Present Study

In the research reported here, we investigated the development of argument representation in a crosslinguistic and developmental study with monolingual English-speaking and monolingual Japanese-speaking children and their mothers. The specific aim of the study was to investigate the relationship between the use of non-linguistic information, such as gestures, and children's referential choices. The children's developmental progress was classified by MLU level, rather than age, in order to investigate the developmental timecourse of the children's learning of argument representation more precisely. The omission and production of sentence arguments was analysed at four different developmental periods between MLU 1.00 and 4.00 (Brown, 1973), and non-linguistic information, such as pointing, reaching, and eye gaze, was systematically analysed. The objective of the gesture analyses was to see whether or not

non-linguistic pragmatic correlates were used in conjunction with sentence arguments. Finally, we investigated the role of input in children's learning of language-specific argument representation. Previous research has shown that, particularly with a language like Japanese where argument omission is permitted and word order varies freely, individual differences might be due to variations in the input children receive (Oshima-Takane, in press).

Using a discourse-pragmatic approach, we hypothesized that children's referential strategies would reflect pragmatic principles of discourse. Specifically, we expected that the children would use non-lexical (null or pronominal) forms to represent given information and lexical forms to represent new information. At earlier stages of language development, on the other hand, we hypothesized that the children would resort to using a non-linguistic referential strategy in communicative attempts if they lacked the appropriate linguistic forms or if they used a referring expression incongruous with discourse-pragmatic expectations. In other words, we expected that the speakers (whether English or Japanese) would use additional non-linguistic pragmatic information to supplement less specific arguments used for referents high in informational value. Although non-linguistic, it was hypothesized that these referential strategies would conform to discourse-pragmatic principles nonetheless. Continuing along the lines of Guerriero, Oshima-Takane, and Kuriyama (2004), we investigated referential choice in English and Japanese because (1) Japanese allows optional omission of subject and object arguments, but English requires both to be overt; and (2) where Japanese allows omission of given information, English dictates pronominalization. However, despite the morphological form that an argument ultimately takes, whether null, pronominal, or

lexical, the same principles are presumed to motivate all forms of argument representation under a discourse-pragmatic approach.

Method

Participants

Two typically-developing English-speaking children (one girl, Nancy, and one boy, Alex)[14] and their mothers and two typically-developing Japanese-speaking children (two boys, Aki and Tai) and their mothers participated in this study. All children received either English-only or Japanese-only input from both parents. All children were first-born, except for the one English-speaking girl, who was second-born. All children came from middle-class families. The data were originally collected for longitudinal projects on language development. The English-speaking children and mothers were observed when the children were 21, 24, 30, and 36 months of age (Oshima-Takane, Goodz, & Derevensky, 1996; Oshima-Takane & Oram, 1991). The Japanese data were obtained from the CHILDES database (MacWhinney, 2000b) and copies of the videotapes were obtained from the contributing researcher (Miyata, 1995; 2000). The children and mothers were observed weekly or biweekly, beginning when the children were approximately 17 months of age until 36 months of age.

[14] "Nancy" and "Alex" are pseudonyms.

Data Collection

The English-speaking children and mothers visited a university playroom that served as an observational room (in Canada); the Japanese-speaking children and mothers were observed in their own homes (in Japan). Each mother-child dyad was individually video-recorded in free-play, naturalistic interaction. Mothers were instructed to play and interact naturally with their children, as they would normally do.

The playroom used by the English-speaking participants was furnished with a table and chairs and equipped with various age-appropriate books and toys, which included puzzles, a set of Megablocks, a farm set, dolls, stuffed bears, Mr. and Mrs. Potato Head, Sesame Street puppets, such as Ernie, Bert, and Kermit, a ricebox, a Fisher Price kitchen set, which contained play food such as hamburgers, hotdogs, mustard, ketchup, assorted vegetables, steak, chicken, eggs, and so forth, and play kitchenware, such as a tea set, pots and pans, plates, drinking glasses, and utensils. The observation room was equipped with a videocamera, television monitor, and microphones suspended from the ceiling. The persons controlling the recording equipment were instructed not to interact with the children and mothers during recordings. Each observational session was recorded for approximately 50 minutes (two 25-minute sessions with a short break in between), except for the final session when the children were 36 months of age, which was recorded for 25 minutes.

The observational sessions for the Japanese-speaking participants took place in their own homes. Using a handheld videocamera, the experimenter recorded the children and mothers while in free-play interaction. The children's books and toys were stored in the family living rooms, which served as playrooms. Because the setting for these

children was their own homes, the toys available for play differed between the two children and from the English-speaking children, but overall, the toys were similar among the four children. Aki's favourite toys included books, stuffed animals, PlayDough, various games and puzzles, building blocks and Lego, play food, assorted toy cars and a garage set, a train set, an airport set, and musical instruments. Tai's favourite toys included books, drawing paper and crayons, puzzles, a dump truck and tractor, a train set, a farm set, building blocks, stuffed animals, such as Donald Duck and Penguin, play food and utensils, and a ball. Also, because the recordings were made in the children's own homes, there was some interaction with the experimenter, or with other family members such as the children's father, grandparents, or a younger sibling, but most of the time the children and mothers interacted with each other. Observational sessions were recorded for approximately 60 minutes for Aki and 40 minutes for Tai.

Transcription and Coding

The children's MLU at each observational session was used to classify individual observational sessions into one of four developmental periods, as indicated below (Barner, Guerriero, & Oshima-Takane, 2001):

Period I (PI) = MLU 1.00 – 1.99 Period II (PII) = MLU 2.00 – 2.99

Period III (PIII) = MLU 3.00 – 3.99 Period IV (PIV) = MLU 4.00 +

Language samples that corresponded to each of the four periods were selected and subjected to intensive coding and analysis. It should be noted that although the same MLU breakdown was used for classifying English and Japanese sessions, English and

Japanese MLU are not directly comparable because of the typological differences between the two languages.

All speech produced by the children and their mothers during the entire recorded observational sessions was transcribed according to the CHAT (English) and JCHAT (Japanese) transcription systems (MacWhinney, 2000a; Oshima-Takane, MacWhinney, Sirai, Miyata, & Naka, 1998). Trained research assistants who were native speakers of English or of Japanese transcribed the sessions from the videotapes. The Japanese sessions were further transcribed to include additional input utterances that had been lacking from the original transcripts downloaded from the CHILDES database. To ensure reliability of the transcriptions, all transcripts were verified for accuracy with the videotapes by a second set of trained research assistants, also native speakers of English or of Japanese. Any disagreements regarding a transcription between the original and second transcriber were verified by a third trained research assistant. If an agreement could not be reached, the CHAT/JCHAT convention was used to mark the utterance as a doubtful transcription.

To obtain a comparable number of utterances for the four children, the first 80 consecutive utterances containing verbs were coded for analysis at each developmental period. For the mothers, coding was restricted to utterances occurring within the first 15 continuous minutes at each developmental period. For some linguistic periods, two or three consecutive observational sessions were combined in order to obtain a sample of 80 utterances containing verbs. A sample of 80 coded utterances was obtained for all children, except for one English-speaking child (Nancy) who produced only 26 utterances containing verbs at Period I. The English data were coded by the first author. Trained

native speakers of Japanese coded the Japanese data. Coding guidelines and criteria are given in Guerriero, Oshima-Takane, Cooper, Kayama, and Ono (2001).

The unit of analysis was the verbal clause. All transitive and intransitive verbs were analysed. Sentence types included declaratives, exclamatives, imperatives (the null subject of an imperative was not coded), tag questions (the argument appearing in the tag was not coded), and *yes-no* interrogatives. Coding was restricted to nominal referring expressions only, whether these appeared as null, pronominal, or lexical arguments. Our decision to code nominal arguments only was based on the difficulty and subjectiveness of operationally defining and quantifying activities and events, in particular for the coding of referential status (see Du Bois, 1987 for a similar argument). All subject and direct object arguments fitting the above criteria were coded for (1) morphological form, (2) referential status (i.e., mention), and (3) non-linguistic pragmatic correlate. Each of these included sub-categories, described below.

First, each argument's morphological form was coded for whether it was a (i) null form (e.g., Ø see the kitty?), (ii) pronominal form (e.g., you see the kitty?), or (iii) lexical form (e.g., the girl sees the kitty). Second, each argument's referential status was coded for whether the argument represented (i) a "given" referent, if the referent was previously mentioned in the preceding 20 clauses (old mention) or (ii) a "new" referent, if the referent was introduced into discourse for the first time (new mention) or if it was mentioned more than 20 clauses previously. First- and second-person pronouns were always coded as "given." Referents present in the situational context, but never referred to, were always coded as "new" upon first mention. Criteria for the coding of referential status followed Chafe (1987) and Du Bois (1987). Third, each argument representing a

referent that was coded as being "present" in the situational context was further coded for whether it was (1) accompanied by a non-linguistic pragmatic correlate or (2) not accompanied by a non-linguistic pragmatic correlate. Types of non-linguistic pragmatic correlates included pointing, touching, reaching, moving, making a head motion, or purposeful gaze direction toward a referent by the speaker at time of mention or pointing, touching, reaching, moving, making a head motion, or purposeful gaze direction toward a referent by the addressee at time of mention.

The following exclusionary criteria based on guidelines given in Guerriero et al. (2001) were applied: Only arguments appearing within main or matrix clauses were coded. Arguments appearing in relative or embedded clauses, subordinate clauses, *to*-infinitive clauses, or nominal *-ing* participle clauses were not coded, though arguments appearing within two compound clauses were. Arguments referring to activities or events (e.g., I went fishing), including deverbal nouns (e.g., give a kiss to Kermie), were excluded from analysis. All arguments of copular verbs (e.g., the dolly is pretty), complex transitive sentences (e.g., I saw her leave the room), ditransitive verbs (e.g., Mommy put the bear in the box), or of *wh*-questions (e.g., who ate the cookie?) were not considered for analysis. Copular verbs were excluded because these verbs are often omitted in Japanese, which is also a grammatical option. This decision came about because (1) the identification and coding of arguments of omitted verbs becomes subjective; and (2) to maximize possibilities for comparison between the English and Japanese data. Also excluded from analysis were subject arguments of verbs appearing in the imperative form (e.g., Ø bring the ball here!). These were excluded by default because omitting the subject of an imperative verb is not an option for speakers, but is

grammatically expected. Finally, only arguments appearing in spontaneous speech were coded; all those appearing in memorized or routine phrases, constituting unanalyzable chunks, such as social routines (e.g., *excuse me*, *thank you*), songs, poems, and nursery rhymes, were excluded.

To determine the reliability of the coding scheme, two transcripts from each of the English and Japanese samples were chosen at random and independently coded by a second set of trained research assistants, also native speakers of English or of Japanese. The reliability coders followed the same guidelines as the original coders and were instructed not to consult or discuss the coding with each other. Proportions of agreement were obtained separately for children and mothers and for English and Japanese. Mean percentages of agreement between the original coders and the reliability coders were 80% for the English-speaking children, 91% for the English-speaking mothers, 66% for the Japanese-speaking children, and 81% for the Japanese-speaking mothers.[15]

Measures and Data Analysis

The CLAN programs (MacWhinney, 2000a) were employed to obtain measures of general language development for each child, such as mean length of utterance (MLU) and total number of intelligible utterances. These measures are based on the entire session

[15] Although the proportion of agreement for the Japanese-speaking children was somewhat lower than that of the other three groups, this should not be taken to mean that the coding scheme was unreliable. We included as disagreements those arguments that were incorrectly coded by either the original coder or the reliability coder (e.g., incorrectly coding a lexical argument as "null"). This conservative method thus underestimates the proportion of agreement. In addition, the higher proportions of agreement obtained for the Japanese-speaking mothers and English-speaking mothers and children indicate that the coding scheme was reliable overall.

for that period. For the data analyses, CLAN programs were used to obtain frequencies of all combinations of argument form (null, pronominal, lexical) and referential status (given, new) for each child and each mother in each language group, resulting in a total of six frequency measures for each individual speaker (i.e., given null, given pronominal, given lexical, new null, new pronominal, and new lexical arguments) at each MLU period. CLAN programs were also employed to obtain frequencies of all combinations of argument form, referential status, and non-linguistic pragmatic correlate (present, absent). These were obtained for each child and each mother at each MLU period.

Results

General Language Measures

The children's ages and general language measures at each developmental period are given in Table 1 (English-speaking children) and Table 2 (Japanese-speaking children). The children's MLUs progressed from Period I to IV, indicating that their linguistic skills developed normally. However, there were individual differences. As shown in Table 1, Nancy's MLU lingered within the same period at each of the four ages, whereas Alex's fluctuated between periods, in particular between Periods I and II at 1;10.28 and 2;3.1. The lower MLU at that age was used for assigning the transcripts to a developmental period. Tai's language development (see Table 2) progressed considerably faster than Aki's. In fact, data are only available for the first three periods for Aki because he did not reach Period IV by the end of his observation period.

TABLES 1 AND 2 ABOUT HERE

Argument Omission and Production

Our analyses of interest were, first, whether the referential status of discourse referents influenced the type of argument form used and, second, whether these patterns were used consistently across developmental periods. To investigate this, we analysed the children's and mothers' use of null, pronominal, and lexical arguments in reference to given versus new information at each developmental period. A child is considered to be sensitive to language-specific discourse-pragmatic strategies if she showed the following patterns: (1) Appropriate reference towards new information requires more frequent use of lexical forms relative to null and pronominal forms, regardless of language. (2a) Appropriate reference towards given information requires more frequent use of pronominal forms relative to null and lexical forms for an English-speaking child. (2b) In contrast, appropriate reference towards given information requires more frequent use of null forms relative to pronominal and lexical forms for a Japanese-speaking child. The proper (language-specific) representation of an argument in syntax requires that the child have mastered the referencing of both given and new information; one cannot be mastered without the other.

The frequencies for Alex and his mother (top) and Nancy and her mother (bottom) are given in Figure 1. At Period I, Alex tended to use null arguments over pronominal or lexical ones in reference to given information. Between Periods I and IV, however, null arguments gradually gave way to pronominal arguments. By Period II,

given referents were referred to with mainly pronominal arguments, though a good number of null arguments were still being used. At Period III, Alex seemed to have mastered the referencing of given information for English: His use of pronominal arguments outnumbered null and lexical ones combined. In regards to new information, Alex used more non-lexical (null and pronominal arguments combined) than lexical arguments at Period I, the majority of which were pronominal forms. By Period II, however, Alex was appropriately using mainly lexical arguments in reference to new information. He continued to use pronominal arguments, but the use of null arguments gradually declined, until no null forms were used at Period IV. From Period II to IV, lexical arguments always outnumbered non-lexical arguments. Similar to Alex, Nancy used mainly null arguments in reference to given information at Period I (see bottom of Figure 1). At Period II, however, the majority of given referents were pronominalized, and her use of pronominal forms continued to increase from Periods II through IV, while null forms decreased. From Period II onwards, Nancy's use of pronominal forms always outnumbered null and lexical forms combined, indicating mastery of the referencing of given information. In regards to new information, from as early as Period I, Nancy appropriately used lexical forms, always outnumbering non-lexical forms. She used null forms only twice; once at Period I and a second time at Period III.

FIGURE 1 ABOUT HERE

The children showed similar patterns to their mothers, though only at later periods of language development: Period III for Alex and Period II for Nancy. Both Alex's and

Nancy's mothers conformed to an English-specific discourse-pragmatic strategy for the referencing of given information from Periods I through IV, whereby given referents were pronominalized. Their use of pronominal forms always outnumbered null and lexical ones combined. In regards to new information, Alex's mother consistently used mainly lexical arguments, though she occasionally used pronominal forms. Likewise for Nancy's mother, new referents were mainly lexicalized, except at Period III, where there was slightly more use of pronominal, rather than lexical, forms. Neither mother ever used null arguments in reference to new information. Overall, the two mothers' referential strategies were consistent with discourse-pragmatic principles.

Aki's and his mother's data (top) and Tai's and his mother's data (bottom) are presented in Figure 2. Aki used predominantly null arguments in reference to given information, always outnumbering pronominal and lexical arguments combined; and this pattern remained the same from Periods I through III. Since Japanese is a null argument language, this is the correct pattern. In regards to new information, Aki tended to use mainly null arguments at Period I. At Periods II and III, he used mainly lexical forms, although non-lexical forms (null and pronominal forms combined) outnumbered lexical ones at Period II; and at Period III, lexical and non-lexical forms were used equally often. Even though Aki was properly referencing given information, his pattern with respect to the referencing of new information suggests that he had not mastered the full spectrum of argument representation under a discourse-pragmatic account even by Period III. Similar to Aki, Tai tended to use predominantly null arguments for referencing given information from Periods I through IV; an appropriate pattern given the null argument option in Japanese. In regards to new information, Tai tended to use mainly null forms at Periods I

and II. Lexical forms were rarely used. At Period III, he used an equal number of null, pronominal, and lexical arguments, whereby the number of non-lexical arguments outnumbered the number of lexical ones. Finally, by Period IV, lexical arguments outnumbered non-lexical ones for the referencing of new information, though only slightly, but suggesting mastery of argument representation skills.

FIGURE 2 ABOUT HERE

Aki's mother conformed to a Japanese-specific discourse-pragmatic strategy in the referencing of given information from Periods I through III: The majority of given referents were represented with a null argument, outnumbering pronominal and lexical arguments combined. In regards to new information, however, Aki's mother showed a pattern somewhat inconsistent with discourse-pragmatic expectations. Except for Period II, where lexical arguments outnumbered non-lexical ones, she used equal numbers of lexical and non-lexical arguments at Periods I and III. Unlike the English-speaking mothers, she never ceased to use null arguments in reference to new information. Similar to Aki's mother, Tai's mother used null arguments in reference to given information from Periods I through IV. She used more non-lexical than lexical arguments in reference to new information at Periods I and II, the majority of which were null forms. At Period III, lexical forms outnumbered non-lexical ones; but at Period IV, non-lexical arguments once again outnumbered lexical ones in the referencing of new information, the majority of which were null forms. In comparison to the English-speaking mothers, the Japanese-

speaking mothers' referential strategies tended to be less consistent overall, and thus less in accordance with discourse-pragmatic expectations.

Interestingly, however, the Japanese-speaking children's patterns were quite similar to their mothers' patterns. This is particularly evident in the referencing of new information. For instance, Tai, whose mother tended to use more null than lexical arguments to reference new information, similarly used more null than lexical arguments in reference to new information. When at Period III Tai's mother used more lexical than non-lexical arguments, Tai showed a similar tendency at Period IV, where lexical arguments outnumbered non-lexical ones. On the other hand, Aki, whose mother tended to use more lexical than null arguments in reference to new information, showed a similar pattern in which he tended to use more lexical than null arguments in reference to new information, especially at Periods II and III. These findings suggest the influence of input in children's learning of argument representation. In the next section, we look at the referencing of new information in more detail.

Use of Non-Linguistic Pragmatic Correlates

In order to test our hypothesis that children would use a non-linguistic referential strategy early on in development if they lacked the appropriate referential expression or if they used a referring expression incongruous with discourse-pragmatic expectations, we specifically analysed the children's and mothers' non-linguistic pragmatic correlates used in conjunction with non-lexical arguments in reference to new information. All null and pronominal arguments were combined into a non-lexical category. Data are presented separately by individual child and her/his mother.

Proportions of non-linguistic pragmatic correlates that are present versus absent for Alex and his mother (top) and Nancy and her mother (bottom) are given in Figure 3. As can be seen in the figure, Alex almost always supplemented non-lexical forms in reference to new information with additional non-linguistic correlates. He used additional non-linguistic information at least 80% of the time. From Period III onwards, Alex supplemented non-lexical uses with additional information 100% of the time. Nancy supplemented non-lexical forms in reference to new information with non-linguistic pragmatic correlates more frequently than Alex did, at least 91% of the time. She began supplementing non-lexical forms in reference to new information from as early as Period I. Similarly to her son, Alex's mother almost always supplemented non-lexical forms in reference to new information with additional non-linguistic correlates. She used additional non-linguistic information at least 80% of the time. Nancy's mother supplemented non-lexical forms in reference to new information with non-linguistic pragmatic correlates at least 83% of the time. Both mothers used additional non-linguistic information from Period II onwards 100% of the time (Alex's mother never produced non-lexical forms in reference to new information at Period IV).

FIGURE 3 ABOUT HERE

Proportions of non-linguistic pragmatic correlates that are present versus absent are shown in Figure 4 for Aki and his mother (top) and Tai and his mother (bottom). A quick glance at the two figures makes it immediately clear that the referential strategies used by the Japanese speakers are quite different from those of the English speakers. Aki

used non-linguistic pragmatic correlates to supplement non-lexical references to new information almost equally as often as he did not, ranging between 50% and 67% of the time. Tai used non-linguistic pragmatic information most of the time, though not always, from 73% to 83% of the time. Aki's mother very infrequently used non-linguistic pragmatic correlates at Periods I and II of Aki's linguistic development. Her usage ranged from 0% to 40% of the time. At Aki's Period III, on the other hand, she supplemented non-lexical forms in reference to new information with non-linguistic information 100% of the time. Tai's mother was somewhat more consistent. She used supplemental non-linguistic information from 64% to 100% of the time.

FIGURE 4 ABOUT HERE

Whether English-speaking or Japanese-speaking, gestures, such as pointing to or touching the referent, were found to be the most common type of non-linguistic pragmatic correlates used by the children to clarify new references made with non-lexical forms. This is true of the children's mothers as well. Several overall patterns emerge: The English-speaking children and mothers used additional non-linguistic pragmatic correlates to clarify references to new information made with less specific argument forms, consistent with discourse-pragmatic expectations, and secondly, the children's patterns were very similar to their mothers' patterns. The Japanese-speaking children and mothers also used additional non-linguistic pragmatic correlates to clarify references to new information made with less specific argument forms, but they showed more variable patterns than their English-speaking counterparts. However, regardless of this variability,

a similar finding to that of the English-speaking children emerges: The Japanese-speaking children's patterns were very similar to their mothers' patterns.

Discussion

The purpose of this study was to investigate the development of language-specific argument representation and its relation, if any, to the use of non-linguistic pragmatic correlates, such as gestures. We had speculated that children would use a non-linguistic referential strategy if they lacked the appropriate referential expression or if the expression chosen was inconsistent with discourse-pragmatic expectations. We also addressed some of the outstanding issues reported in Guerriero, Oshima-Takane and Kuriyama (2004). Specifically, rather than grouping children by age, we grouped children's linguistic abilities by MLU level. The age grouping in Guerriero, Oshima-Takane, and Kuriyama's study could have contributed to their inconclusive and unexpected findings. Finally, we examined the children's input, since this too could be a factor contributing to variability in children's productions.

Our study supports the earlier finding of Guerriero, Oshima-Takane, and Kuriyama (2004) that English-speaking and Japanese-speaking children do not show sensitivity to discourse-pragmatic principles of communication at early stages of linguistic development. The English-speaking children, Alex and Nancy, were not appropriately referencing given and new information at Period I (between MLU 1.00 and 1.99). Mastery of discourse-pragmatic skills began to surface between Periods II and III (between MLU 2.00 and 3.99). Consistent with an English-specific discourse-pragmatic strategy, Nancy seemed to have mastered pronominalization of given referents and

lexicalization of new referents from Period II. Alex, on the other hand, did not master appropriate discourse-pragmatic strategies until Period III. While he showed lexicalization of new referents from Period II, he did not master pronominalization of given referents until Period III. Such individual variation is not new in child language, and confirms similar individual variation obtained by Guerriero, Oshima-Takane, and Kuriyama. In contrast to that study, however, the present one suggests that English-speaking children master discourse-pragmatic strategies for language-specific argument representation even before 36 months of age (3;0). Alex, who showed mastery at Period III, was 2;7.17 and Nancy, who showed mastery at Period II, was even younger; she was 2;0.8. As for the children's mothers, the two were very similar to each other. Across the four developmental periods, Alex's and Nancy's mothers pronominalized given referents, in accordance with an English-specific strategy. New referents were lexicalized, likewise in accordance with discourse-pragmatic principles. This was the case across all four developmental periods, except for Nancy's mother at Period III, where she used more pronominal than lexical references. These findings are consistent with those of Guerriero, Oshima-Takane, and Kuriyama, who similarly found that the English-speaking mothers in their study tended to use lexical and non-lexical arguments in accordance with (English-specific) discourse-pragmatic strategies. Overall, the children's referential patterns closely resembled their mothers' patterns, and such similarity became manifest even before Period IV (MLU 4.00 and above).

The Japanese-speaking children, Aki and Tai, appropriately used null arguments to refer to given information, in accordance with a Japanese-specific discourse-pragmatic strategy. Unlike the English-speaking children, the Japanese-speaking children used null

arguments in reference to given information from as early as Period I, and this remained consistent over time. Since Japanese allows null arguments, this finding was expected. However, the children's referential choices for the expression of new information were unexpected. At Period I, the use of non-lexical arguments in reference to new information, a greater number of which were null arguments, outnumbered lexical arguments. Non-lexical arguments outnumbered lexical ones until Period III for Aki and Period IV for Tai, where lexical arguments were used equally often (Aki) or slightly more frequently (Tai) than non-lexical ones. The same patterns appear in the Japanese-speaking mothers' data. While the referencing of given information was made mostly with null forms, the referencing of new information was somewhat inconsistent between the use of non-lexical (mainly null) and lexical arguments. This is especially evident with Tai's mother, for whom non-lexical arguments in reference to new information, particularly the use of null forms, seemed to be the rule rather than the exception. Two overall conclusions can be made with regards to the Japanese-speaking children and mothers: First, the children's referential patterns closely resembled their mothers' patterns, and secondly, the children's and mothers' referential choices seemed to contradict discourse-pragmatic principles for the referencing of new information.

Analyses of the children's and mothers' use of non-linguistic pragmatic correlates suggest that, here too, English and Japanese speakers differ from each other. Our prediction, that a non-linguistic pragmatic strategy would be used when an appropriate referential expression was lacking or inconsistent with discourse-pragmatic principles, was supported with respect to the English-speaking children and mothers. Results revealed that the mothers supplemented most non-lexical references to new information

with gestures. A strikingly similar pattern was obtained with the English-speaking children: They supplemented non-specific references to new information, such as references made with a null or a pronominal argument, with additional non-linguistic information, particularly pointing and touching gestures. This behaviour was observed from as early as Period I, and indicates that the children were in fact sensitive to the higher informational value of new mentions, even though they did not use a more specific lexical form. Together, the results suggest that at some point between MLU 1.00 and 1.99 English-learning children are sensitive to the informational status of discourse referents and are capable of coordinating their linguistic and non-linguistic communicative strategies in accordance with pragmatic principles of discourse. These findings are consistent with previous studies investigating the referential nature of children's gestures (e.g., O'Neill & Topolovec, 2001; Tomasello, Anselmi, & Farrar, 1985; Wilcox & Howse, 1982).

The situation is somewhat different with the Japanese speakers. The children's and mothers' use of non-linguistic pragmatic correlates was found to be more variable than that of the English speakers'. Null and pronominal arguments in reference to new information were supplemented with non-linguistic information, particularly pointing and touching gestures, though many were not. Unlike the results obtained by Guerriero, Oshima-Takane and Kuriyama (2004), our results do not support the hypothesis that Japanese speakers supplement less specific references to new information with non-linguistic pragmatic correlates. Moreover, we obtained a high degree of individual variability. For instance, Tai's mother used additional non-linguistic correlates more consistently than Aki's mother. Similar to the finding with the English speakers,

strikingly similar patterns were observed between the Japanese-speaking children and their mothers. Aki, whose mother showed an inconsistent pattern, used non-linguistic information equally as often as he did not; whereas Tai, whose mother used non-linguistic information at least three-fourths of the time, similarly used non-linguistic information at least three-fourths of the time. Together, the linguistic and non-linguistic referential strategies used by the Japanese speakers, especially those used when referencing new information, suggest that the children and mothers might not have been sensitive to (presumably universal) discourse-pragmatic strategies.

One possible explanation for this seemingly peculiar use of null arguments is that the Japanese-speaking children and mothers were observed in their own homes, in which the contextual situation was familiar and well known to both. In such a situation, a referent mentioned for the first time (consequently defined and coded as "new") was not, in essence, "new" to the participants, but simply new to the discourse conversation due to its being mentioned for the first time at that particular observational session. As such, the children and mothers might have felt it unnecessary to lexicalize an argument representing a referent that had already been experienced, perhaps even several times, in past conversations. However, a similar finding of the use of null arguments in reference to new information was made by Guerriero, Oshima-Takane, and Kuriyama (2004), and their observational setting was a playroom at a university in which the toys and situational context were unfamiliar to the participants. Thus, although this might seem a likely explanation on the surface, it becomes inadequate when scrutinized a little further.

A more likely interpretation is that the Japanese-speaking children and mothers possessed shared contextual knowledge and thus modified their linguistic (and non-

linguistic) strategies accordingly. We suspect that because of this shared contextual knowledge between the children and their mothers, as well as the familiarity between the speakers, null references to new information were allowed as pragmatically acceptable (Clancy, 1986a). Typically, when reference to a new item is made, principles of discourse communication necessitate a lexical mention, however the situational context would allow a non-lexical mention if the speaker and listener possess shared background knowledge. Others have reported on young children's ability and skill at taking the prior knowledge and experiences possessed by the listener into account during referential communication and to vary their speech as a function of such listener characteristics (e.g., O'Neill & Topolovec, 2001; Sonnenschein, 1986). This would explain how the Japanese speakers were able to make use of null arguments in reference to new information whether in a familiar home setting (as in this study) or an unfamiliar university playroom (as in Guerriero, Oshima-Takane, and Kuriyama's (2004) study). In fact, the use of null arguments in reference to new information has similarly been observed with Inuktitut-speaking children (Allen, 2000; Skarabela & Allen, 2002). Perhaps null argument languages might allow a more lax set of discourse-pragmatic principles, in particular during mother-child interactions, likely due to the informal nature of such exchanges. Moreover, the Japanese coders reported that miscommunication between the children and their mothers rarely, if ever, occurred; indicating that the use of null arguments for the referencing of new information, even if used without any accompanying non-linguistic pragmatic information, was an effective strategy.

While the idea that Japanese might employ more lax referential strategies requires further study, it nevertheless cannot be said that the Japanese-speaking children were

using deviant or unconventional discourse-pragmatic strategies, since similar strategies were used by their own mothers. It is also not the case that the Japanese-speaking children might not have known what was "new" in the discourse situation. Recent research by Tomasello and Haberl (2003) indicates that children as young as 12 months of age know what is "new" for other persons. In their experimental study, German-speaking children were able to select which of three objects was new for their parent, while also old ("given") for the children (the parent was out of the room when the new object was presented to the child).

The inconsistency observed in the Japanese speakers' use of gestures lends itself to yet another possible interpretation. Studies investigating the relationship between gesture and speech production point to a unified and integrated communication system, whereby the structure and content of gestures parallel the structural and semantic properties co-expressed in speech (Butcher & Goldin-Meadow, 2000; Mayberry & Jaques, 2000; also McNeill, 1992). For instance, gestures and speech show semantic coherence (gestures are combined with meaningful and related speech) and temporal synchrony (gestures are produced in synchrony with speech). In other words, gestures encode information that is simultaneously also expressed in speech. However, if a language does not verbally encode a particular referential concept, the corresponding gestural expression is similarly not produced (Kita & Özyürek, 2003). Thus, in a null argument language such as Japanese, which does not require overt mention of arguments, a gestural expression might be omitted when a sentence argument is similarly omitted.

Likewise, because overt arguments are required in English, they are gestured as well.[16] If this interpretation is correct, it suggests the existence of crosslinguistic variability in the link between gestures and syntax, which is expected given the two languages' differing typologies (Kita & Özyürek, 2003). At this point in time, however, the interpretation is merely speculation. Whether gestures are linked to Japanese null arguments in the same way that they are linked to English overt arguments requires further investigation.

What the present research does show, however, is that the close similarity of referential patterns between the Japanese-speaking children and their mothers suggests that children's referential choices are strongly motivated by parental input strategies. This is true of the English-learning children as well, for whom referential choices observed in parental input were the least variable. It seems that children exposed to more consistent input tend to make more consistent referential choices. The Japanese-learning children, on the other hand, were exposed to somewhat inconsistent input, especially in regards to the referencing of new information, and such inconsistency was reproduced by the children. The individual variability observed between the two Japanese-speaking dyads furthermore suggests that the children were not merely adapting to the language of their environment, in the sense that English requires pronouns, whereas Japanese allows null arguments. Rather, the Japanese-speaking dyads' idiosyncratic linguistic and non-linguistic patterns support the proposal that the children learned the patterns from input.

Others have similarly reported that parental input has a strong influence on children's linguistic patterns. For instance, verb frequency in input is a good predictor of

---

[16] We thank Rachel Mayberry for suggesting this interpretation to us.

verb frequency in children's speech at a later time (Naigles & Hoff-Ginsberg, 1995; 1998). Children learning null argument languages such as Mandarin Chinese and Korean, for whom verbs might be more salient due to frequent argument omission, tend to show a verb spurt (a verb bias) in early vocabulary, in contrast to children learning overt argument languages, who show a noun spurt (Choi & Gopnik, 1995; Tardif, Shatz, & Naigles, 1997). It has also been shown that parental use of an intransitive versus a transitive frame with verbs of mixed transitivity is a good predictor of the specific frames used by their children with those same verbs (Theakston, Lieven, Pine, & Rowland, 2001). More striking evidence is reported in Miyata, Oshima-Takane, and Nisisawa (2004) who found that some Japanese-learning children showed a noun over a verb bias due to individual parental input patterns: Children whose mothers used a noun-focused speech style displayed a noun-biased vocabulary, whereas children whose mothers used nouns and verbs equally often produced a more balanced vocabulary. These studies reinforce our proposal that parental input is a strong influence on children's referential patterns and, furthermore, that individual variability is accountable by variation in parental input. Although the role played by input needs further investigation, the present study has shown that young English-speaking and Japanese-speaking children learn discourse-pragmatic strategies, both linguistic and non-linguistic patterns, via parental input.

Table 1
General Language Measures: English-Speaking Children

Period	Age	MLU	Total Intelligible Utterances	Length of Session
Alex				
PI	1;10.28	1.945	145	25"
		2.269	175	25"
PII	2;3.1	3.417	115	25"
		2.746	134	25"
PIII	2;7.17	3.574	204	25"
PIV	3;1.16	4.871	224	25"
Nancy				
PI	1;8.29	1.664	107	25"
		1.737	190	25"
PII	2;0.8	2.754	65	28"
		2.404	183	27"
PIII	2;6.12	3.205	151	29"
		3.491	165	26"
PIV	3;0.4	4.087	218	25"

Note. Periods I and II for Alex were determined based on the lower MLU at that age.

Table 2
General Language Measures: Japanese-Speaking Children

Period	Age	MLU	Total Intelligible Utterances	Length of Session
Aki				
PI	2;3.4	1.584	507	61"
	2;3.12	1.724	486	60"
PII	2;6.29	2.786	346	59"
	2;7.5	2.723	300	63"
	2;7.12	2.818	413	60"
PIII	2;11.25	3.343	233	35"
	3;0.0	3.363	394	60"
Tai				
PI	1;5.20	1.514	274	40"
	1;5.27	1.591	287	47"
	1;6.4	1.288	264	30"
PII	1;11.1	2.467	462	41"
PIII	2;6.3	3.486	440	41"
PIV	2;9.9	4.779	401	42"

Note. MLU and number of utterances were obtained from Miyata (1995, 2000).

Figure Captions

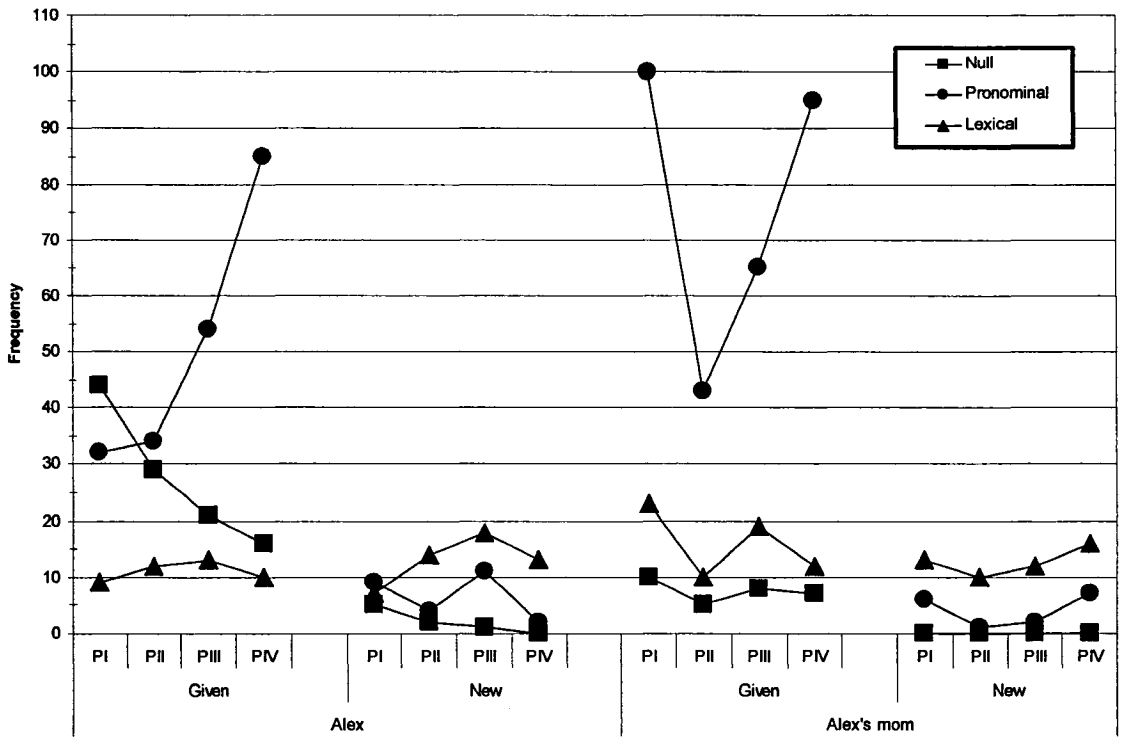
Figure 1. Frequencies of null, pronominal, and lexical arguments in reference to given and new information for Alex and his mother (top figure) and Nancy and her mother (bottom figure) at Periods I, II, III, and IV.

Figure 2. Frequencies of null, pronominal, and lexical arguments in reference to given and new information for Aki and his mother (top figure) and Tai and his mother (bottom figure) at Periods I, II, and III. (Aki's data does not reach Period IV.)

Figure 3. Proportions of non-linguistic pragmatic correlates that are present versus absent for non-lexical (null and pronominal) arguments used in reference to new information for Alex and his mother (top figure) and Nancy and her mother (bottom figure) at Periods I, II, III, and IV. (Alex's mother never used non-lexical forms in reference to new information at Period IV.)

Figure 4. Proportions of non-linguistic pragmatic correlates that are present versus absent for non-lexical (null and pronominal) arguments used in reference to new information for Aki and his mother (top figure) and Tai and his mother (bottom figure) at Periods I, II, III, and IV.

Figure 1


Null
Pronominal
Lexical
Frequency
PI
PII
PIII
PIV
Given
New
Alex
Alex's mom

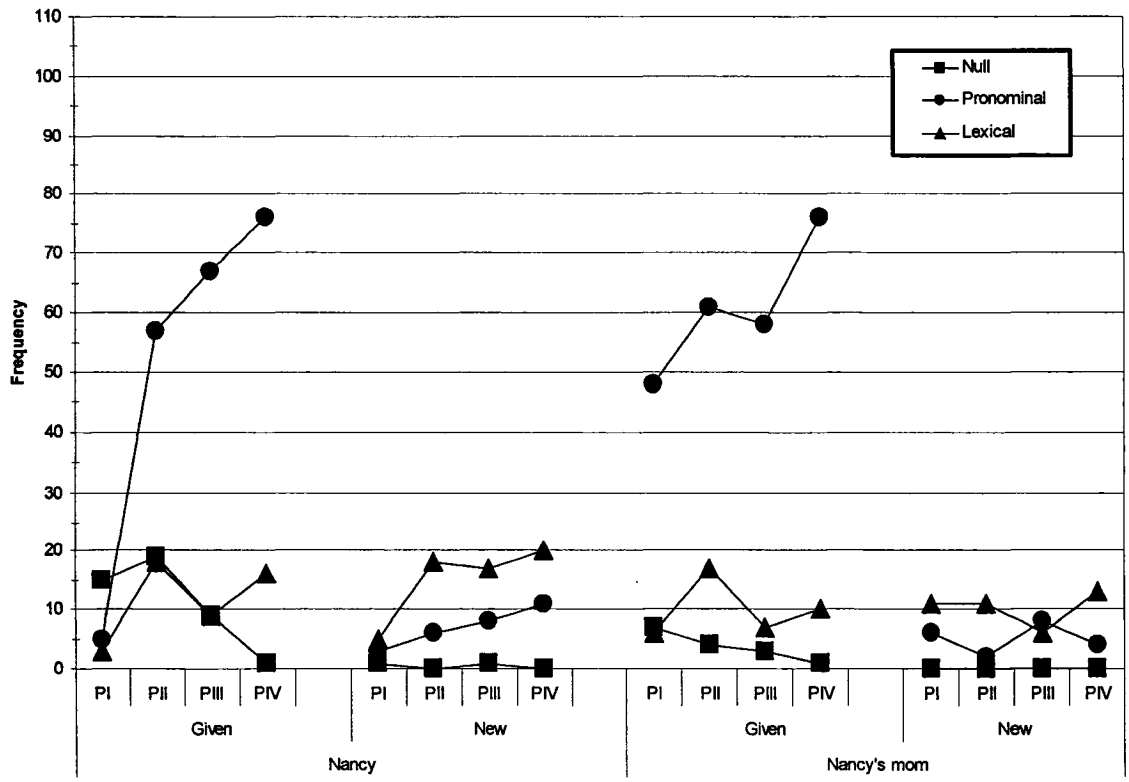

Null
Pronominal
Lexical
Frequency
PI
PII
PIII
PIV
Given
New
Nancy
Nancy's mom

Figure 2

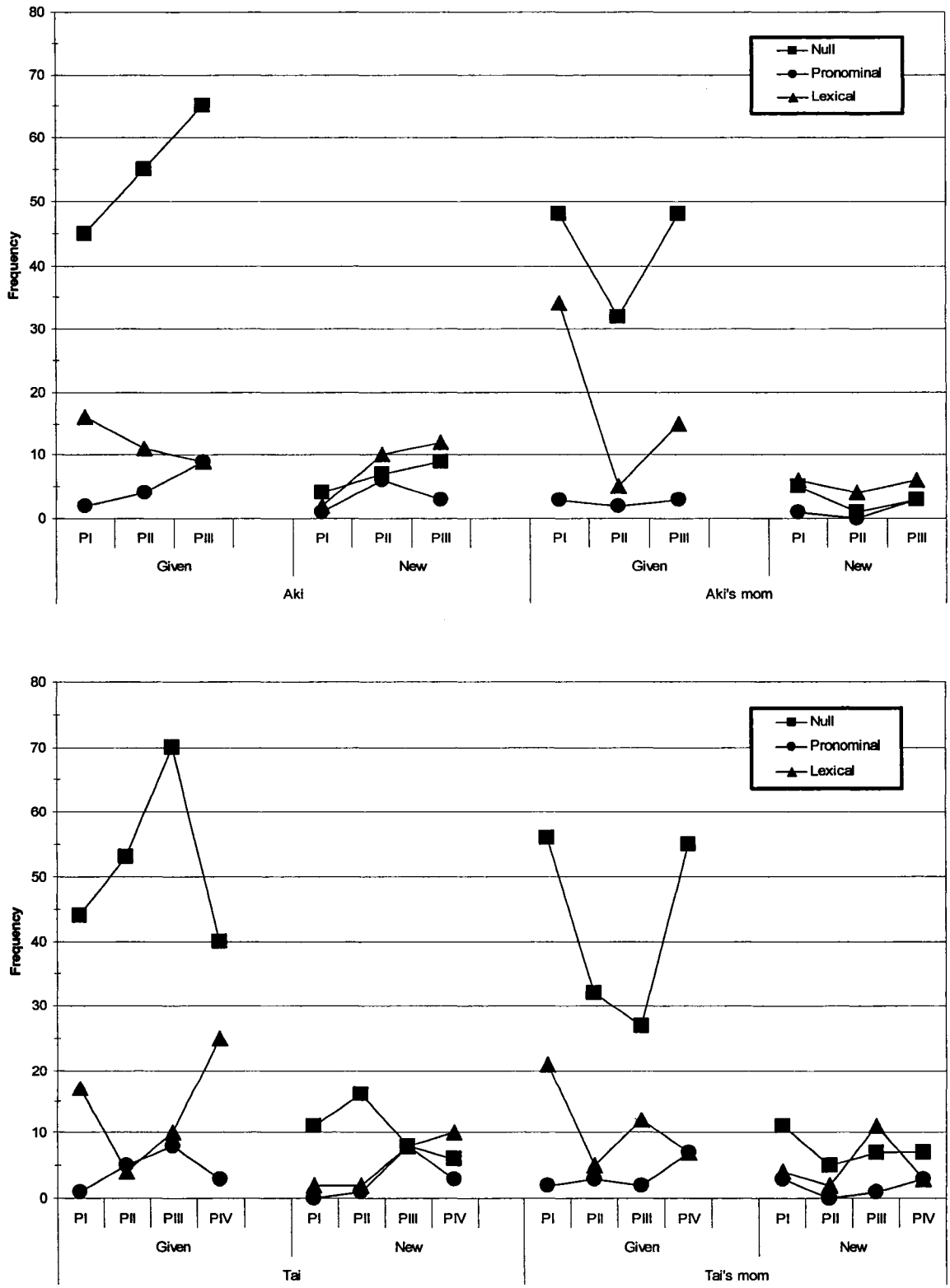

Null
Pronominal
Lexical
Frequency
80
70
60
50
40
30
20
10
0
PI
PII
PIII
PIV
Given
New
Aki
Aki's mom
Tai
Tai's mom

Figure 3

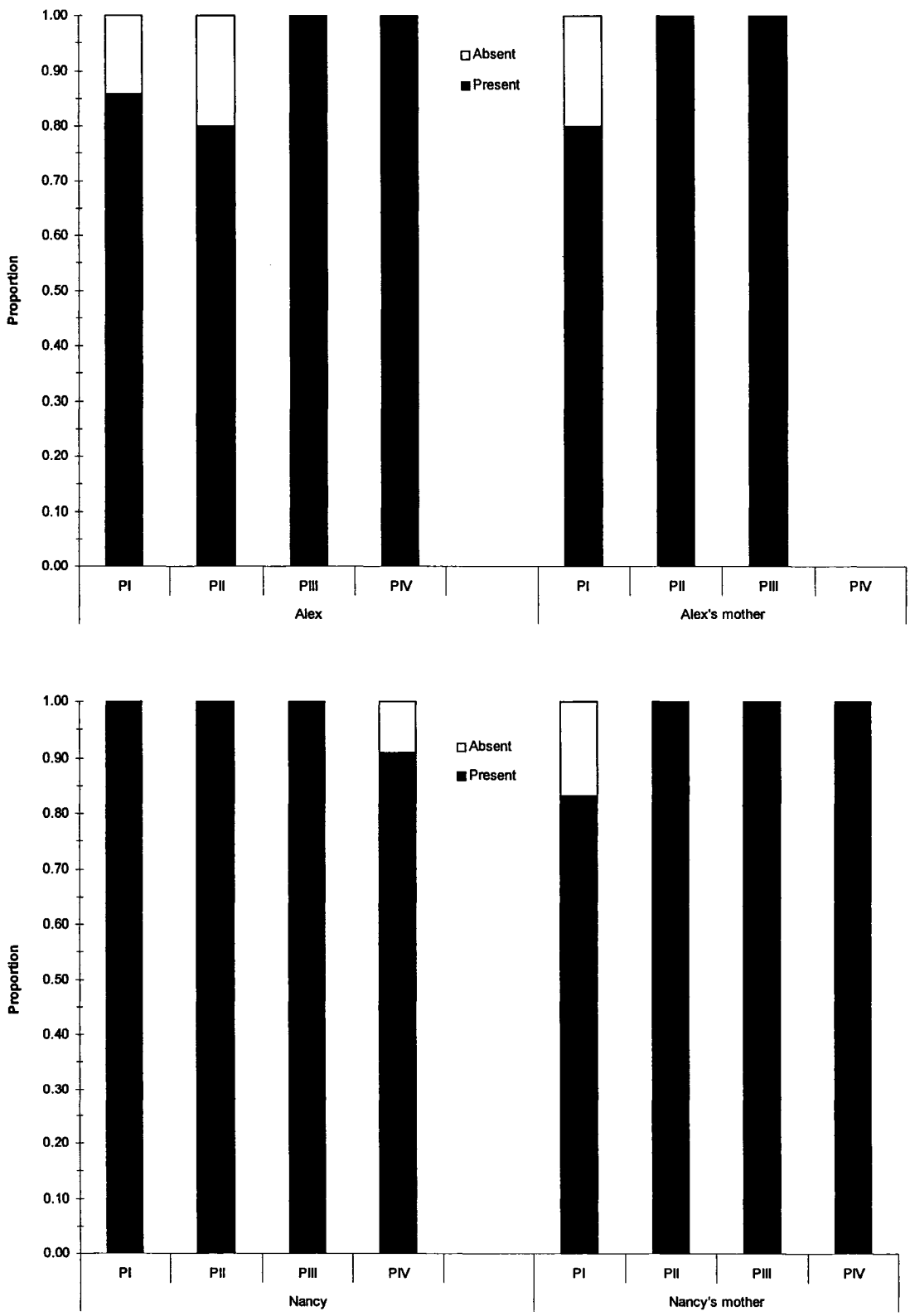
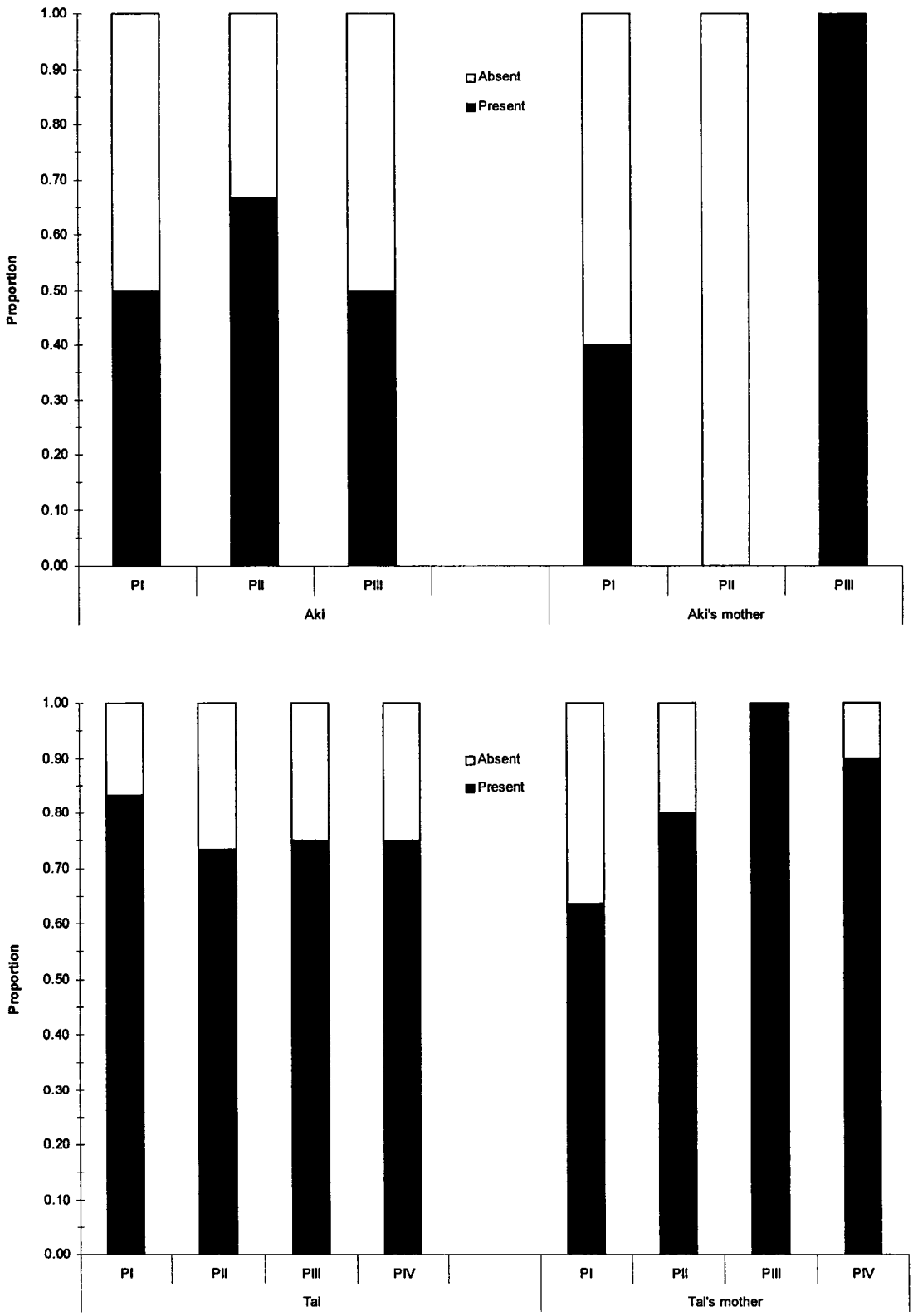

1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
Proportion
Absent
Present
PI
PII
PIII
PIV
Alex
Alex's mother
Nancy
Nancy's mother

Figure 4


1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
Proportion
Absent
Present
PI
PII
PIII
Aki
PI
PII
PIII
Aki's mother
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
Proportion
Absent
Present
PI
PII
PIII
PIV
Tai
PI
PII
PIII
PIV
Tai's mother

CHAPTER 4

Transition to Manuscript

The third manuscript in this series investigated children's sensitivity to Preferred Argument Structure (PAS; Du Bois, 1987) in a crosslinguistic and developmental study of English-speaking and Japanese-speaking children at 21 and 36 months of age. Previous studies have shown that children are sensitive to the strategies underlying Preferred Argument Structure (e.g., Allen & Schröder, 2003; Clancy, 2003); however, the question of whether PAS strategies uphold crosslinguistically and developmentally remains unresolved. Previous research showed that children tend to avoid using more than one new or lexical argument per transitive clause and tend to avoid casting new or lexical arguments as transitive subjects. To further investigate these patterns, the study presented here also analysed arguments which were simultaneously both new *and* lexical, which have not been analysed before. The children's input is also analysed in order to investigate the relationship between language-universal and language-specific discourse-pragmatic strategies in children's learning of PAS patterns.

Running Head: UNIVERSAL AND LANGUAGE-SPECIFIC FACTORS IN ARGUMENT DISTRIBUTION

MANUSCRIPT IN PREPARATION FOR SUBMISSION

The Syntactic Distribution of Sentence Arguments in Child Language: The Role of Universal and Language-Specific Factors

A M Sonia Guerriero and Yuriko Oshima-Takane

McGill University

Yoko Kuriyama

International Christian University

This research was supported by a Social Sciences and Humanities Research Council of Canada internal grant to the first author and a Natural Sciences and Engineering Research Council of Canada grant to the second author. We thank Amy Cooper, Yuhko Kayama, and Alyssa Ono for their help in transcribing and coding the Japanese data and Fred Genesee and Tom Shultz for helpful comments. We especially thank the parents and children who participated in the study.

Address for correspondence: A M Sonia Guerriero, Department of Psychology, McGill University, 1205 Docteur Penfield Avenue, Montréal, Québec, H3A 1B1, CANADA. Email: sonia.guerriero@mail.mcgill.ca.

Abstract

The relationship between argument form (null, pronominal, lexical), referential status (given, new) and syntactic role (subject, object) was investigated in monolingual English-speaking and Japanese-speaking children and their mothers when the children were 21 and 36 months of age. Results support Du Bois's (1987) Preferred Argument Structure (PAS) constraints: Children and mothers tended to avoid using more than one new or lexical argument per transitive clause and tended to avoid casting new or lexical arguments as transitive subjects. These findings are the first to show that PAS strategies are upheld crosslinguistically and developmentally. Universal PAS constraints seemed to be in place from early stages of language development, but language-specific referential choices (such as omission versus pronominalization of a given argument) developed over time. Findings suggest that the learning of argument representation and its distribution in syntax results from the interaction between language-universal discourse-pragmatic principles and language-specific linguistic strategies, learned via parental input.

The Syntactic Distribution of Sentence Arguments in Child Language: The Role of Universal and Language-Specific Factors

Introduction

Recent research investigating the relationship between syntax and discourse suggests the presence of systematic discourse-pragmatic strategies that are presumed to motivate the morphological representation of discourse referents: Referents that are "new" to the discourse (i.e., mentioned for the first time) tend to be represented by full lexical NPs, whereas referents that are "given" (i.e., previously mentioned) tend to be represented by pronominal forms, or if the language permits it, omitted altogether. It has been argued that this given/new dimension of discourse referents reflects the underlying cognitive, or "activation," state of the concepts represented in memory (Chafe, 1987; 1994; 1996). Given referents are concepts said to be active in a person's consciousness and, as such, reside in working memory, whereas new referents are those which are either inactive in a person's consciousness and, thus, stored in long-term memory, or those which have never entered consciousness before. Because given referents are presumed to lie in working memory for both interlocutors – having been mentioned previously in the discourse, they are known to both speaker and listener – their verbal expression can be reduced to a pronominal or null argument. By contrast, new referents – new to the discourse and, as such, retrieved from long-term memory – are typically expressed as lexical arguments as the speaker's intention is to be as explicit as possible for the listener to unambiguously identify the newly introduced referent. Thus, the morphological representation of sentence arguments suggests an underlying systematic patterning

between grammar and the referential status (i.e., the recency of mention) of discourse referents. This systematic patterning is proposed to be true of all languages and has been observed in English (Chafe, 1987; 1994; Gundel, Hedberg, & Zacharski, 1993; Kärkkäinen, 1996), Seneca (Chafe, 1987; 1994), Sakapultek (Du Bois, 1987), Hebrew, German, Brazilian Portuguese, (cited in Du Bois, 1987), French, Spanish (Ashby & Bentivoglio, 1993), Japanese, Mandarin Chinese, Russian (Gundel, Hedberg, & Zacharski, 1993), among others, suggesting a link between language and cognition.

The link between language and cognition is further illuminated by investigations of the distribution of sentence arguments within syntax, in which it has been proposed that the surface form of utterances is motivated by the activation state of discourse referents. Chafe (1987; 1994) proposed a "light starting point constraint" (also called a "light subject constraint") in which the grammatical subject of a sentence is ascribed a unique and special status because it acts as a "starting point" to which new information is then attached. Sentences conform to a light starting point because speakers tend to choose subjects expressing referents that are already active ("given") in the discourse context. Subjects thus carry a light information load; and, as such, given referents make appropriate starting points for sentences because they represent less costly referents. Since these referents are already active in a person's consciousness, less mental effort is required for the referent to be identified. On the other hand, new referents are costly in comparison to given ones because new concepts are recalled from long-term memory and, as such, exert greater cognitive effort. Chafe's analyses of English and Seneca (a Native American language spoken on reservations in western New York State) provide support for the "light starting point" constraint. He found that the majority of subject

arguments expressed given information and these were typically represented by pronominal forms.

Another of Chafe's (1987; 1994) observations, namely, the amount of information that can be communicated within a single clause, further suggests a link between language and cognition. Chafe's analyses revealed that speakers of English and Seneca rarely expressed more than one previously inactive ("new") concept within a single utterance. Moreover, because of the restriction inherent in the "light starting point" constraint, the newly introduced information tended to appear in the predicate. Thus, the distribution of information within an utterance was typically that of given information in the subject and new information in the predicate. This finding was interpreted as an upper limit on the amount of information that can be fully active in the mind at any one time (i.e., within one utterance). Chafe suggested that such a linguistic limit results from the cognitive basis of a linguistic clause and labelled it the "one new argument constraint." The constraint is assumed to reflect the presence of an underlying cognitive limitation not only on the part of the speaker, but that of the listener as well, and is likely related to the limited capacities of working memory. Together, the "light starting point" and "one new argument" constraints shape the surface form of discourse utterances and suggest ways in which the mind handles the flow of information through consciousness and language.

The surface distribution of sentence arguments has also been investigated by Du Bois (1987), who analysed Sakapultek narrative discourse by adult speakers (Sakapultek is a Mayan language spoken in highland Guatemala). Du Bois examined the morphological form (null, pronominal, lexical), grammatical role (subject, object), and referential status (given, new; what Du Bois termed the "information" status of referents)

of sentence arguments and uncovered that certain patterns of information and grammar tended to co-occur consistently. These patterns, which he termed "Preferred Argument Structure," reflected statistical tendencies underlying the distribution of sentence arguments in syntax and discourse. New and lexical arguments tended to appear as subjects of intransitive verbs or objects of transitive verbs, but very rarely as subjects of transitive verbs. The subject role of transitive verbs tended to be reserved for given and non-lexical (i.e., null or pronominal) arguments. Secondly, clauses, whether transitive or intransitive, tented to contain either zero or just one lexical or new argument. The occurrence of two lexical or two new arguments within a single (transitive) clause was extremely rare, suggesting a cognitive limit on producing more than one lexical or new argument within a single utterance. Du Bois interpreted these tendencies in terms of constraints on the syntactic role and quantity of sentence arguments on two dimensions, pragmatic and grammatical, and like Chafe, proposed that PAS patterning would underlie all languages. PAS has, in fact, been observed in several typologically distinct languages, such as English (Kärkkäinen, 1996), Hebrew, German, Brazilian Portuguese (as cited in Du Bois, 1987), French, Spanish (Ashby & Bentivoglio, 1993), and Japanese (Matsumoto, 2000), furthermore suggesting that PAS constraints might be universal. The dimensions and constraints of PAS as described above are outlined in Table 1.

TABLE 1 ABOUT HERE

Recently, PAS patterning has been investigated in child language as well. Clancy (1993; 1997; 2003) analysed argument choice and the syntactic distribution of sentence

arguments in two children learning Korean. The children, aged 1;8 and 1;10 at the start, were observed longitudinally for a period of one year. Allen (1997; 2000) and Allen and Schröder (2003) conducted similar analyses with four children learning Inuktitut who were observed for a period of nine months. At the start of the study two children were aged 2;6 and two others were aged 2;0 and 2;10. Analyses from both languages produced similar results: The children showed a tendency to realize as overt those arguments representing informative referents. Informativeness was characterized into features of pragmatic prominence, of which newness was one feature. By contrast, arguments representing uninformative referents, such as those characterized as given, tended to be omitted. Moreover, the Inuktitut-speaking and Korean-speaking children's language conformed to PAS patterning: New and lexical arguments tended to be cast as subjects of intransitive verbs or objects of transitive verbs, but very infrequently as subjects of transitive verbs, indicating sensitivity to the syntactic role constraints. The children furthermore showed a tendency to produce only one new or lexical argument per clause, indicating sensitivity to the quantity constraints. Together, these findings indicate that children, like adults, show sensitivity to the statistical tendencies underlying Preferred Argument Structure.

However, it is unclear whether such strategies are upheld developmentally. Although Allen's (1997; 2000), Allen and Schröder's (2003), and Clancy's (1993; 1997; 2003) studies reported longitudinal child language data, the data were not analysed developmentally. It is thus unclear how early in development PAS strategies emerge and, secondly, whether the patterns remain consistent throughout development. It is also unclear whether childhood PAS generalizes crosslinguistically. To date, PAS has been

observed in several typologically distinct adult languages, but only in Korean and Inuktitut child language, both of which are null argument languages. Evidence of PAS strategies in an overt argument child language would further support their universality.

A related issue is the role played by parental input in children's learning of discourse-pragmatic strategies, such as PAS. Although input has been touted as misleading or inconsistent (e.g., Valian, 1990), its role in acquisition is nevertheless necessary; otherwise children would not learn essential language-specific grammatical options. A recent study investigating children's referential choices showed that children acquiring English learned to pronominalize given information – information that was previously omitted at an earlier stage of language development – from patterns observed in maternal input. Likewise, children acquiring Japanese, who heard omission of given information in their input, continued to omit given information (Guerriero & Oshima-Takane, 2004; also Guerriero, Oshima-Takane, & Kuriyama, 2004). Undoubtedly, children can only learn such language-specific grammatical properties, such as whether the target language is an overt or a null argument one, via parental input. Several other studies have shown that the influence of parental input on children's language output is highly significant and likely responsible for the variation seen in individual children's language patterns. For instance, Miyata, Oshima-Takane, and Nisisawa (2004; also Oshima-Takane, in press) reported that Japanese-speaking children whose mothers used a noun-focused speech style displayed a noun-biased vocabulary, whereas the children whose mothers used nouns and verbs equally often produced a more balanced vocabulary. The authors suggested that the children's individual parental input patterns accounted for why some children showed a noun over a verb bias whereas others did not.

Guerriero and Oshima-Takane (2004) showed that individual input patterns also influenced children's non-linguistic referential strategies, such as the use of a gesture towards the referent in question: A Japanese-speaking child whose mother used non-linguistic information at least three-fourths of the time similarly used non-linguistic information at least three-fourths of the time; whereas another child whose mother showed an inconsistent pattern used non-linguistic information equally as often as he did not. These findings suggest that children are quite sensitive to parental input patterns. Many researchers have made similar suggestions (Choi & Gopnik, 1995; Naigles & Hoff-Ginsberg, 1995; 1998; Rowland, Pine, Lieven, & Theakston, 2003; Tardif, Shatz, & Naigles, 1997; Theakston, Lieven, Pine, & Rowland, 2001; 2004).

The Present Study

In the present study, the morphological representation and distributional patterns of arguments in syntax and discourse (what is known as "Preferred Argument Structure") in early child language was examined. We conducted a crosslinguistic comparison of two typologically different languages, Japanese and English. The use of Japanese and English as comparison languages was not an arbitrary choice, since one of our goals was to investigate the distributional patterns of sentence arguments in both null and overt argument child languages. Only by employing the same methodology crosslinguistically could the differential contributions of language-specific and language-universal patterns be observed. No study to date has attempted a similar design. In order to look at developmental patterns, we analysed the children's language at two time points in development, 21 and 36 months of age. Finally, we investigated the relationship between

parental input patterns and children's learning of discourse-pragmatic strategies. Although previous research has observed the use of PAS strategies in adult speech to children (e.g., Clancy, 1993), none has specifically examined whether input plays a role in children's learning of such patterns.

The findings reported here are based on analyses of sentence arguments in the children's and mothers' spontaneous language productions with respect to the argument's morphological form (null, pronominal, lexical), referential status (given, new), and syntactic role (transitive subject, transitive object, intransitive subject). We hypothesized that the same universal discourse-pragmatic strategies, namely, the patterns underlying Preferred Argument Structure, would account for argument choice and distribution in both English and Japanese. Based on previous PAS findings, we expected both groups of children and mothers to avoid using more than one new or more than one lexical argument per clause (the One New Argument and One Lexical Argument constraints; see Table 1). Secondly, we expected both English-speaking and Japanese-speaking children and mothers to show a tendency to avoid casting new and lexical arguments as subjects of transitive verbs (the Given A and Non-Lexical A constraints). Prior investigations have examined these constraints by analysing arguments that are independently either new *or* lexical, and although the findings suggest similar restrictions with respect to arguments that are simultaneously both new *and* lexical, this has yet to be empirically tested. Thus, in addition to our analyses of the Given A and Non-Lexical A constraints, we also analysed arguments that were simultaneously both new and lexical. We expected these language-universal patterns to be evident from as early as 21 months of age. However, language-specific referential options, such as whether a language allowed

argument omission as a grammatical option, were hypothesized to be learned from patterns observed in parental input. We thus expected the children to reproduce respective language-specific input patterns by 36 months of age; that is, the English-speaking children to show pronominalization of given information and the Japanese-speaking children to show omission. The mothers, by contrast, were expected to provide their children with consistent patterns from one age period to the next. Before we describe the methodology, we briefly review the relevant grammatical features of English and Japanese.

The Structures of Japanese and English

We chose English and Japanese because Japanese allows optional omission of either subject or object arguments, whereas English requires both to be overt. For example, except for the sentence in 2a, the English sentences illustrated in 2b, 2c, and 2d below are ungrammatical. On the other hand, all four of the Japanese equivalents are grammatical:[17]

(2) a. *Hanako-ga ringo-o kat-ta.*

Hanako-NOM apple-ACC buy-PAST.

b. *Ø ringo-o kat-ta.*

Ø apple-ACC buy-PAST.

[17] The following abbreviations are used in glosses: Ø = omitted argument; ACC = accusative case marker; NOM = nominative case marker; PAST = past tense marker.

c. *Hanako-ga Ø kat-ta.*

Hanako-NOM Ø buy-PAST.

d. *Ø Ø kat-ta.*

Ø Ø buy-PAST.

This grammatical difference between English and Japanese makes them interesting cases for crosslinguistic analyses. Unlike other null argument languages, such as Italian in which recovery of a null subject pronoun is possible due to rich inflectional markings on the verb, Japanese does not mark verbal agreement. In other words, Japanese verbs do not give any information about the identity (e.g., person, number, or gender) of the referent of an omitted argument. Case is marked via the use of case markers. However, an omitted argument cannot be recovered via its case marker because the case marker is also omitted when the argument is omitted. Recovery of the identity of a null argument in Japanese is typically made via the discourse or pragmatic context (Clancy, 1986b; Wang, Lillo-Martin, Best, & Levitt, 1992). English, in contrast, is a language in which argument omission is not permitted. Case is marked via word order, though some pronouns have nominative and accusative forms. There are but a few overt agreement inflections on the verb, such as the third-person singular *-s* and the past-tense *-ed*. Table 2 summarizes the major features of the two languages:

TABLE 2 ABOUT HERE

Method

Participants

Six typically-developing English-speaking and six typically-developing Japanese-speaking girls and their mothers participated in the study. All children received either English-only or Japanese-only input from both parents. All children were first-born and came from middle-class families. The data were originally collected for longitudinal projects on language development. The English-speaking girls and mothers were observed when the children were 21, 24, 30, and 36 months of age (Oshima-Takane, Goodz, & Derevensky, 1996; Oshima-Takane & Oram, 1991) and the Japanese-speaking girls and mothers were observed when the children were 12, 21, 30, 36, and 48 months of age (Kuriyama, 1999). The results reported here are based on data collected when the children were 21 and 36 months of age. At the 21-month observational sessions, the English-speaking children ranged in age from 1;08.29 to 1;10.14; and at the 36-month observations, they ranged from 2;11.30 to 3;01.05. The Japanese-speaking children ranged in age from 1;08.23 to 1;10.17 at the 21-month observational sessions and from 2;11.29 to 3;00.24 at the 36-month observations.

Data Collection

The children and their mothers visited a university playroom that served as an observational room. English-speaking participants were video-recorded in Canada; Japanese-speaking participants in Japan. Each mother-child dyad was individually video-recorded in free-play, naturalistic interaction. Mothers were instructed to play and interact naturally with their children, as they would at home. The playrooms were

furnished with a table and chairs and equipped with various age-appropriate books and toys, such as puzzles, a set of Megablocks, a farm set, dolls, stuffed bears, Mr. and Mrs. Potato Head, Sesame Street puppets, such as Ernie, Bert, and Kermit, a ricebox, a Fisher Price kitchen set, which contained play food such as hamburgers, hotdogs, mustard, ketchup, assorted vegetables, steak, chicken, eggs, and so forth, and play kitchenware, such as a tea set, pots and pans, plates, drinking glasses, and utensils. The observation rooms were equipped with videocameras, television monitors, and microphones suspended from the ceiling. The persons controlling the recording equipment were instructed to not interact with the children and mothers during recordings.

Play sessions with the English-speaking dyads were recorded for approximately 50 minutes when the children were 21 months of age (two 25-minute sessions with a short break in between) and approximately 25 minutes when they were 36 months of age. For the Japanese-speaking dyads, the play sessions were recorded for approximately 15 minutes when the children were 21 months of age and approximately 25 minutes when they were 36 months of age. For the purpose of this study, the first 15 minutes of each session were used for analysis.

Transcription and Coding

Twelve play sessions (6 English, 6 Japanese) at each of the two age periods (21 months, 36 months) resulted in a total of 24 sessions for analysis. All speech produced by the children and their mothers during the entire recorded play sessions was transcribed according to the CHAT (English) and JCHAT (Japanese) transcription systems (MacWhinney, 2000a; Oshima-Takane, 1998). Trained research assistants who were

native speakers of English or of Japanese transcribed the sessions from the videotapes. To ensure reliability of the transcriptions, all transcripts were verified for accuracy with the videotapes by a second set of trained research assistants. Any disagreements regarding a transcription between the original and second transcriber were verified by the original transcriber or a third trained research assistant. If an agreement could not be reached, the CHAT/JCHAT convention was used to mark the utterance as a doubtful transcription.

Children's and mothers' language in the first 15 continuous minutes of each of the 24 transcripts was coded and analysed. The English transcripts were coded by the first author. Trained native speakers of Japanese coded the Japanese sample. Coding guidelines and criteria are given in Guerriero, Oshima-Takane, Cooper, Kayama, and Ono (2001).

The unit of analysis is the verbal clause. All transitive and intransitive verbs were analysed. Sentence types included declaratives, exclamatives, tag questions (the argument appearing in the tag was not coded), and *yes-no* interrogatives. Coding was restricted to nominal referring expressions only, whether these appeared as null, pronominal, or lexical arguments. Our decision to code nominal arguments only was based on the difficulty and subjectiveness of operationally defining and quantifying activities and events, in particular for the coding of information status (see Du Bois, 1987 for a similar argument). All subject and direct object arguments fitting the above criteria were coded for (1) morphological form, (2) referential status (i.e., mention), and (3) syntactic role. First, each argument was coded for whether it was represented as a (i) null form (e.g., Ø see the kitty?), (ii) pronominal form (e.g., you see the kitty?), or (iii) lexical form (e.g., the girl sees the kitty). Second, each argument was coded for whether the referent it

represented was (i) "given," if the referent was previously mentioned in the preceding 20 clauses or (ii) "new," if the referent was introduced into discourse for the first time or if it was mentioned more than 20 clauses previously. First- and second-person pronouns were always coded as "given." Referents present in the situational context, but never referred to, were always coded as "new" upon first mention. Criteria for the coding of referential status followed Chafe (1987) and Du Bois (1987). Finally, each argument was coded for whether it appeared in the (i) subject position of a transitive clause, (ii) object position of a transitive clause, or (iii) subject position of an intransitive clause.

The following exclusionary criteria based on guidelines given in Guerriero et al. (2001) were applied: Only arguments appearing within main or matrix clauses were coded. Arguments appearing in relative or embedded clauses, subordinate clauses, *to*-infinitive clauses, or nominal *-ing* participle clauses were not coded, though arguments appearing within two compound clauses were. Arguments appearing within sentences referring to activities or events (e.g., I went fishing), including deverbal nouns (e.g., give a kiss to Kermie), were excluded from analysis. All arguments of copular verbs (e.g., the dolly is pretty), complex transitive sentences (e.g., I saw her leave the room), ditransitive verbs (e.g., Mommy put the bear in the box), or of *wh*-questions (e.g., who ate the cookie?) were not considered for analysis. Note that the entire sentence was excluded in such cases. Copular verbs were excluded from analysis because these verbs are often omitted in Japanese, which is also a grammatical option. This decision came about because (1) the identification and coding of arguments of omitted verbs becomes subjective; and (2) to maximize possibilities for comparison between the English and Japanese data. Also excluded from analysis were subject and object arguments of verbs

appearing in the imperative form (e.g., Ø bring the ball here!). These sentences were excluded by default because omitting the subject of an imperative verb is not an option for speakers, but is grammatically expected. Finally, only arguments appearing in spontaneous speech were coded; all those appearing in memorized or routine phrases, constituting unanalyzable chunks, such as social routines (e.g., *excuse me*, *thank you*), songs, poems, and nursery rhymes, were excluded.

To determine the reliability of the coding scheme, two out of six transcripts from each of the English and Japanese samples (33.33%) were chosen at random and independently coded by a second set of trained research assistants, also native speakers of English or Japanese. The reliability coders followed the same guidelines as the original coders and were instructed to not consult and/or discuss the coding with each other. Proportions of agreement were obtained separately for children and mothers and for English and Japanese. Mean percentages of agreement between the original coders and the reliability coders were 85% for the English-speaking children, 88% for the English-speaking mothers, 77% for the Japanese-speaking children, and 78% for the Japanese-speaking mothers.

Measures and Data Analysis

The CLAN programs (MacWhinney, 2000a) were employed to obtain general language measures for each child, such as mean length of utterance (MLU) and total number of intelligible utterances. These measures were based on the first 15 minutes of continuous speech in each transcript. For the data analyses, CLAN programs were used to obtain frequencies of transitive and intransitive clauses and the number (zero, one, or

two) of arguments each contained. CLAN programs were also employed to obtain frequencies of all combinations of syntactic role (subject of transitive verb, object of transitive verb, subject of intransitive verb), argument form (null, pronominal, lexical) and referential status (given, new). Frequencies were obtained for each speaker at each age group. However, rather than directly comparing frequencies between groups, we report patterns within each language group. This is because the purpose of the study was to investigate whether the patterns observed are consistent with universal and language-specific discourse-pragmatic patterns, rather than investigating crosslinguistic differences in frequencies between groups.

Results

General Language Measures

The six English-speaking children produced an average of 106.83 intelligible utterances at 21 months of age (ranging from 45 to 144) and an average of 116.83 intelligible utterances at 36 months of age (ranging from 54 to 197). Their mean MLUs were 1.44 (ranging from 1.09 to 2.25) and 3.62 (ranging from 2.90 to 4.15) at 21 and 36 months of age, respectively. The Japanese-speaking children were not talkative at 21 months and only produced an average 22.33 intelligible utterances (ranging from 8 to 57). However, by 36 months of age, they produced an average of 92.50 intelligible utterances (ranging from 42 to 120). Their mean MLUs were 1.46 (ranging from 1.07 to 1.91) and 3.61 (ranging from 2.86 to 4.29) at 21 and 36 months of age, respectively. Because only the first 15 minutes of speech was used, some children's MLUs are based on less than 100 utterances.

Frequency of Verbs and Arguments

Table 3 presents group frequencies of the total number of verbs (clauses) and arguments included for analysis in this paper. Data are separated between children and mothers and grouped by language and age. Results show that English and Japanese speakers showed different patterns with respect to verb transitivity. The English-speaking children and mothers tended to use more transitive over intransitive verbs overall. All mothers showed this pattern consistently across age periods. This tendency was similarly observed with the children across age periods, except for one child at 21 months who produced more intransitive over transitive verbs.

TABLE 3 ABOUT HERE

In contrast to the English speakers, the Japanese-speaking mothers tended to use more intransitive over transitive verbs overall. The mothers showed this pattern across age periods, except for one mother at 36 months who showed a more balanced pattern, using transitive and intransitive verbs equally often. Although the Japanese-speaking children showed an overall greater use of intransitive over transitive verbs at 21 months of age, they generally spoke very little at this age, thus the results are somewhat inconclusive. Of the six children, only four produced verbs; and of these four children, three produced verbs only once the entire session (two children produced one intransitive verb each and one child produced one transitive verb; the only transitive verb produced at this age). By 36 months, the Japanese-speaking children's verb productions increased considerably, though as a group they used an almost equal proportion of transitive and

intransitive verbs. Three children used intransitives more frequently and three used transitives more frequently. This is different from the pattern observed with the Japanese-speaking mothers, who showed a clear bias in using intransitive verbs more frequently over transitive ones. We return to this discrepancy later.

Quantity Constraints

Our first analyses of interest concern the quantity constraints of Preferred Argument Structure. These involve the number of new and lexical arguments that can be produced within a transitive clause. Although the quantity constraints do not apply to intransitive clauses (only transitive verbs can take more than one argument), we have included the results with respect to both types of verbs for overall comparison purposes.

One New Argument Constraint. The children's and mothers' mean proportions and frequencies (and standard deviations) of zero, one, and two new arguments produced within transitive and intransitive clauses are given in Table 4. We expected both groups of children and mothers to avoid using more than one new argument with transitive clauses, as per the One New Argument Constraint.

Regardless of age period, the English-speaking mothers rarely produced two new arguments within transitive clauses. A single-mean *t*-test revealed that the mothers' mean frequency of two new arguments was not significantly different from zero at 21 months ($M = 0.17$, $t(5) = 1.00$, $p > 0.05$, one-tailed) nor at 36 months ($M = 1.50$, $t(5) = 1.96$, $p > 0.05$, one-tailed), supporting the One New Argument Constraint. Similar to their mothers, the English-speaking children conformed to the One New Argument Constraint. They

never produced more than one new argument with transitive clauses at 21 months of age and the mean frequency of two new arguments at 36 months ($M = 0.17$) was not significantly different from zero ($t(5) = 1.00, p > 0.05$, one-tailed). Interestingly, neither the mothers nor the children typically used transitive and intransitive clauses containing one new argument. They used transitive clauses containing zero new arguments 41% to 67% of the time. As for intransitive clauses, they used clauses containing zero new arguments most of the time (79%–92%).

TABLE 4 ABOUT HERE

The Japanese-speaking mothers rarely produced transitive clauses containing more than one new argument. Mean frequencies of two new arguments were not significantly different from zero at both age periods ($M = 0.17$ at both 21 and 36 months: $t(5) = 1.00, p > 0.05$, one-tailed), thus supporting the One New Argument Constraint. Similarly, the Japanese-speaking children conformed to the One New Argument Constraint as they never used two new arguments with transitive clauses at 36 months of age. (Data are not given at 21 months because only one child produced one transitive verb at this age). The mothers used transitive clauses containing zero new arguments more frequently than those containing one new argument across age periods (69%–86%), whereas the children used them about equally often (54% versus 46%). As for intransitive clauses, the children used zero new argument clauses most of the time across age periods (83%–97%), but the mothers used them less so (64%–72%).

One Lexical Argument Constraint. The children's and mothers' mean proportions (and standard deviations) of zero, one, and two lexical arguments produced within transitive and intransitive clauses are summarized in Table 5. Based on the One Lexical Argument Constraint, we expected the children and the mothers to avoid using more than one lexical argument per transitive clause.

Across age periods, producing more than one lexical argument per transitive clause was rare for the English-speaking mothers. However, the mean frequency of lexical arguments was significantly different from zero at 21 months ($M = 1.83$, $t(5) = 2.61$, $p < 0.05$), although not at 36 months ($M = 1.00$, $t(5) = 1.58$, $p > 0.05$). Although the mothers' use of two lexical arguments was significantly different from zero at 21 months, the mean proportion of two lexical arguments (7%) was smaller than that of zero (39%) or one lexical argument (54%), thus supporting the One Lexical Argument Constraint. Similar to their mothers, the use of more than one lexical argument was rare for the English-speaking children as well. The mean frequency of two lexical arguments was not significantly different from zero at either age period ($M = 0.17$ at both 21 and 36 months: $t(5) = 1.00$, $p > 0.05$, one-tailed). Neither the mothers nor the children typically used one lexical argument for either transitive or intransitive clauses. In fact, the children's mean proportion of transitive clauses containing one lexical argument was very low (15%) at 21 months of age, but it increased to 49% by 36 months of age. By contrast, the mothers used clauses containing one lexical argument slightly more frequently than those containing zero lexical arguments across age periods (54%–58%). As for intransitive clauses, both the mothers and children used zero lexical arguments most of the time (89%–94%).

TABLE 5 ABOUT HERE

The Japanese-speaking mothers rarely used two lexical arguments with transitive clauses. The mean frequencies of two lexical arguments were not significantly different from zero at both age periods (21 months: $M = 0.50$, $t(5) = 1.46$, $p > 0.05$, one-tailed; 36 months: $M = 1.33$, $t(5) = 1.87$, $p > 0.05$, one-tailed), supporting the One Lexical Argument Constraint. The Japanese-speaking children also conformed to the One Lexical Argument Constraint. The mean frequency of two lexical arguments at 36 months of age was not significantly different from zero ($M = 0.67$, $t(5) = 2.00$, $p > 0.05$, one-tailed). Similar to English-speaking group, the Japanese-speaking mothers and children did not typically use one lexical argument for either transitive or intransitive clauses. They used transitive clauses containing zero or one lexical argument about equally often (41%–50%). As for intransitive clauses, the children produced zero lexical arguments most of the time (83%–85%) across age periods, whereas the mothers produced them less so (65%–69%).

Syntactic Role Constraints

Our second analyses of interest concern the role constraints of Preferred Argument Structure. These involve the distribution of lexical and new arguments across the available grammatical roles for transitive and intransitive clauses. We use the notation as outlined in Du Bois (1987): A = transitive subject; O = transitive object; S = intransitive subject (see Table 1).

Non-Lexical A Constraint. Mean frequencies and proportions (and standard deviations) for the children's and mother's use of non-lexical (null and pronominal combined) and lexical arguments within each of the three grammatical roles are summarized in Table 6 (English speakers) and Table 7 (Japanese speakers). Based on the Non-Lexical A Constraint, we expected the children and mothers to avoid casting lexical arguments in A role. As such, the A position was expected to contain more non-lexical than lexical arguments.

The mean frequencies of lexical A's for the English-speaking mothers were significantly different from zero at 21 months ($M = 2.83$, $t(5) = 3.00$, $p < 0.05$, one-tailed), but not at 36 months ($M = 1.00$, $t(5) = 1.58$, $p > 0.05$, one-tailed). Although the mothers' use of lexical A's was significantly different from zero at 21 months, their use of non-lexical A's far outnumbered lexical ones (90% versus 11%), thus supporting the Non-Lexical A Constraint. Furthermore, a Tukey HSD test revealed that the mothers' use of lexical A's did not differ significantly between 21 and 36 months ($Q(6, 25) = 2.16$, $p > 0.05$, two-tailed), indicating that the pattern remained consistent across the two age periods. The English-speaking children also conformed to the Non-Lexical A constraint. At both age periods, the mean frequencies of lexical A's did not differ significantly from zero (21 months: $M = 0.33$, $t(5) = 1.00$, $p > 0.05$, one-tailed; 36 months: $M = 0.17$, $t(5) = 1.00$, $p > 0.05$, one-tailed). A Tukey HSD test revealed that the children's use of lexical A's did not differ significantly between 21 and 36 months of age ($Q(6, 25) = 1.41$, $p > 0.05$, two-tailed), indicating no change over time.

The English-speaking mothers rarely used lexical arguments in S position as well (from 10% to 12% of the time), although the mean frequencies of lexical S's were

significantly different from zero at both age periods (21 months: $M = 1.50$, $t(5) = 2.24$, $p < 0.05$, one-tailed; 36 months: $M = 1.50$, $t(5) = 2.67$, $p < 0.05$). On the other hand, the mothers tended to use lexical arguments more frequently than non-lexical arguments in O position across age periods (from 58% to 61% of the time) and the mean frequencies of lexical O's were significantly different from zero (21 months: $M = 19.17$, $t(5) = 6.07$, $p < 0.05$, one-tailed; 36 months: $M = 15.83$, $t(5) = 5.86$, $p < 0.05$, one-tailed). There were no significant differences in the mean frequencies of lexical arguments in either S or O positions between 21 and 36 months of age, indicating that the pattern of use of lexical versus non-lexical arguments was consistent across age periods. The mean frequencies of lexical arguments did not differ significantly between A and S positions across age periods (21 months: $Q(6, 25) = 1.69$, $p > 0.05$, two-tailed; 36 months: $Q(6, 25) = -1.25$, $p > 0.05$, two-tailed), though they differed significantly between A and O positions (21 months: $Q(6, 25) = -6.36$, $p < 0.05$, two-tailed; 36 months: $Q(6, 25) = -8.17$, $p < 0.05$, two-tailed) as well as between O and S positions (21 months: $Q(6, 25) = 8.25$, $p < 0.05$, two-tailed; 36 months: $Q(6, 25) = 7.61$, $p < 0.05$). This suggests that the English-speaking mothers tended to avoid using lexical arguments not only in A position but in S position as well.

TABLE 6 ABOUT HERE

Similar to their mothers, the English-speaking children rarely used lexical arguments in either S or O positions at 21 months of age (11%–15%) and the mean frequencies of lexical arguments in both positions did not differ significantly from zero

(lexical S's: $M = 0.33$, $t(5) = 1.58$, $p > 0.05$, one-tailed; lexical O's: $M = 3.00$, $t(5) = 1.15$, $p > 0.05$, one-tailed). However, lexical O's differed significantly from zero at 36 months of age ($M = 11.33$, $t(5) = 4.14$, $p < 0.05$, one-tailed). The children's distribution of lexical arguments over the three syntactic roles is more similar to their mothers' patterns at 36 months of age than at 21 months of age, suggesting some learning over time.

The mean frequency of lexical A's for the Japanese-speaking mothers was not significantly different from zero at 21 months ($M = 3.17$, $t(5) = 1.71$, $p > 0.05$, one-tailed), but was significant at 36 months ($M = 4.00$, $t(5) = 3.16$, $p < 0.05$, one-tailed). Although the use of lexical A's was significantly different from zero at 36 months, non-lexical A's outnumbered lexical ones (76% versus 24%), thus supporting the Non-Lexical A Constraint. A Tukey HSD test revealed that the mean frequencies of lexical A's did not differ significantly between the two age periods ($Q(6, 25) = -0.55$, $p > 0.05$, two-tailed), indicating that the mothers' pattern of use was consistent from one age period to the next. Similar to their mothers, the results for the Japanese-speaking children at 36 months of age conformed to the Non-Lexical A Constraint. The children's use of non-lexical arguments in A position outnumbered that of lexical arguments (77% versus 24%), although the mean frequency of lexical A's was significantly different from zero ($M = 2.83$, $t(5) = 3.58$, $p < 0.05$, one-tailed).

Unlike the English-speaking mothers, the Japanese-speaking mothers preferred non-lexical over lexical arguments not only in A and S positions, but also in O position across age periods (62%–69%), although both lexical S's and lexical O's were significantly different from zero (lexical S's: 21 months: $M = 9.50$, $t(5) = 3.35$, $p < 0.05$, one-tailed; 36 months: $M = 9.83$, $t(5) = 7.27$, $p < 0.05$, one-tailed; lexical O's: 21

months: $M = 2.67$, $t(5) = 3.02$, $p < 0.05$, one-tailed; 36 months: $M = 5.50$, $t(5) = 5.75$, $p < 0.05$, one-tailed). Tukey HSD tests revealed that the mean frequency of lexical A's did not differ significantly from that of lexical S's at 21 months ($Q(6, 25) = -4.21$, $p > 0.05$, two-tailed), but did differ significantly at 36 months ($Q(6, 25) = -6.20$, $p < 0.05$, two-tailed), indicating that at 36 months the mothers avoided using lexical arguments in S position to a lesser extent than in A position. The mean frequency of lexical A's was not significantly different from that of lexical O's at 21 months ($Q(6, 25) = 0.31$, $p > 0.05$, two-tailed), nor at 36 months ($Q(6, 25) = -2.06$, $p > 0.05$, two-tailed). The mean frequencies of lexical arguments in O and S positions did not differ significantly from each other at 21 months ($Q(6, 25) = -3.76$, $p > 0.05$, two-tailed), but did at 36 months ($Q(6, 25) = -7.25$, $p < 0.05$, two-tailed).

TABLE 7 ABOUT HERE

Like their mothers, the Japanese-speaking children preferred to use non-lexical over lexical arguments in S position (15%–17%), although the mean frequency of lexical S's was significantly different from zero only at 36 months (21 months: $M = 1.00$, $t(5) = 1.00$, $p > 0.05$, one-tailed; 36 months: $M = 2.67$, $t(5) = 2.27$, $p < 0.05$, one-tailed). Also similar to the Japanese-speaking mothers, the O position tended to be non-lexical rather than lexical (34%), although the mean frequency of lexical O's was significantly different from zero at 36 months ($M = 3.67$, $t(5) = 5.97$, $p < 0.05$, one-tailed). Tukey HSD tests revealed no significant differences between lexical A's and O's nor between lexical A's and S's, indicating no differences among the three roles (lexical A's vs. lexical O's: $Q(6,$

25) = –1.29, $p > 0.05$, two-tailed; lexical A's vs. lexical S's: $Q(6, 25) = 0.25$, $p > 0.05$, two-tailed).

In sum, although there are language-specific variations in overall frequency, our findings provide support for Du Bois's (1987) constraint that the use of lexical A's is crosslinguistically avoided. Both groups of children and mothers tended to use non-lexical arguments more frequently than lexical ones in A position (the subject role of transitive verbs). Moreover, the children's tendencies closely resembled those of their input. Whereas the English-speaking mothers tended to cast lexical arguments in O position, the Japanese-speaking mothers tended to show a more balanced distribution of lexical arguments across the three grammatical roles. These patterns were similarly observed in their respective children's patterns when the children were 36 months of age.

Given A Constraint. Mean frequencies and proportions (and standard deviations) for the children's and mother's use of given versus new arguments within each of the three syntactic roles are given in Table 8 (English speakers) and Table 9 (Japanese speakers). We expected the children and mothers to avoid casting new arguments in A role, as indicated by the Given A Constraint. Thus, the A role was expected to contain more given arguments than new arguments.

The mean frequency of new A's for the English-speaking mothers was not significantly different from zero at 21 months ($M = 0.33$, $t(5) = 1.58$, $p > 0.05$, one-tailed), though it was significantly different at 36 months ($M = 2.33$, $t(5) = 2.15$, $p < 0.05$, one-tailed). A Tukey HSD test indicated that new A's did not differ significantly between 21 and 36 months ($Q(6, 25) = –3.15$, $p > 0.05$, two-tailed), suggesting no change in use

over time. Overall, however, the A role contained more given than new arguments at both age periods (91%–99%), supporting the Given A Constraint. Like their mothers, the English-speaking children conformed to the Given A Constraint. They never used new arguments in A position at 21 months of age. At 36 months, their use of given A's far outnumbered new A's (95% versus 5%), although the mean frequency of new A's differed significantly from zero ($M = 0.50$, $t(5) = 2.24$, $p < 0.05$, one-tailed).

Both the English-speaking mothers and children disfavoured new arguments in S position. They used given arguments most of the time (79%–92%), although the mothers' mean frequency of new S's was significantly different from zero at both times (21 months: $M = 1.17$, $t(5) = 2.15$, $p < 0.05$, one-tailed; 36 months: $M = 2.33$, $t(5) = 3.50$, $p < 0.05$, one-tailed). Across the three syntactic roles, the majority of new arguments tended to appear in O position and the mean frequencies of new O's differed significantly from zero at both times for both the mothers (21 months: $M = 11.17$, $t(5) = 4.79$, $p < 0.05$, one-tailed; 36 months: $M = 12.67$, $t(5) = 4.54$, $p < 0.05$, one-tailed) and the children (21 months: $M = 3.33$, $t(5) = 2.84$, $p < 0.05$, one-tailed; 36 months: $M = 14.00$, $t(5) = 3.20$, $p < 0.05$, one-tailed). Tukey HSD tests indicated that the mothers' new A's differed significantly from new O's across age periods (21 months: $Q(6, 25) = -6.36$, $p < 0.05$, two-tailed; 36 months: $Q(6, 25) = -5.60$, $p < 0.05$, two-tailed), but not from new S's (21 months: $Q(6, 25) = -1.67$, $p > 0.05$, two-tailed; 36 months: $Q(6, 25) = 0.00$, $p > 0.05$, two-tailed). As for the children, the difference between new A's and new O's was not significant at either 21 ($Q(6, 25) = -4.00$, $p > 0.05$, two-tailed) or 36 months of age ($Q(6, 25) = -4.16$, $p > 0.05$, two-tailed). Mean frequencies between new A's and new S's were

also not significantly different at either age period (21 months: $Q(6, 25) = -1.90, p > 0.05$, two-tailed; 36 months: $Q(6, 25) = -0.59, p > 0.05$, two-tailed).

TABLE 8 ABOUT HERE

The Japanese-speaking mothers used given arguments more frequently than new arguments in A position across age periods (98%). The mothers' mean frequencies of new A's were not significantly different from zero at both ages (21 months: $M = 0.50$, $t(5) = 1.00, p > 0.05$, one-tailed; 36 months: $M = 0.33$, $t(5) = 1.58, p > 0.05$, one-tailed), supporting the Given A Constraint. New A's did not differ significantly between the two age periods ($Q(6, 25) = 0.39, p > 0.05$, two-tailed), indicating consistent use over time. The Japanese-speaking children similarly conformed to the Given A Constraint, as they never used new A's at 36 months of age.

The Japanese-speaking mothers used given arguments the majority of the time in both S and O positions (64%–87%), although the mean frequencies of new S's and new O's were significantly different from zero across age periods (new S's: 21 months: $M = 8.50$, $t(5) = 5.16, p < 0.05$, one-tailed; 36 months: $M = 7.83$, $t(5) = 10.46, p < 0.05$, one-tailed; new O's: 21 months: $M = 2.50$, $t(5) = 2.37, p < 0.05$, one-tailed; 36 months: $M = 2.00$, $t(5) = 4.47, p < 0.05$, one-tailed). Unlike the English-speaking group, the Japanese-speaking mothers tended to cast new arguments in S position more frequently than in O position across age periods. The mean frequencies of new A's differed significantly from new S's at both age periods (21 months: $Q(6, 25) = -8.74, p < 0.05$, two-tailed; 36 months: $Q(6, 25) = -15.76, p < 0.05$, two-tailed). New A's did not differ significantly

from new O's at 21 months ($Q(6, 25) = -4.13, p > 0.05$, two-tailed), but differed significantly at 36 months ($Q(6, 25) = -7.05, p < 0.05$, two-tailed).

TABLE 9 ABOUT HERE

Similar to their mothers, the Japanese-speaking children used given arguments in both S and O positions the majority of the time (54%–97%). However, unlike their mothers, the children located more new arguments in O position (46%) than S position (17%) and the mean frequency of new O's was significantly different from zero ($M = 4.83, t(5) = 6.87, p < 0.05$, one-tailed). The mean frequency of new A's was significantly different from new O's ($Q(6, 25) = -9.69, p < 0.05$, two-tailed), but not from new S's ($Q(6, 25) = -3.37, p > 0.06$, two-tailed) at 36 months of age. The mean frequency of new S's was not significantly different from zero at 21 months ($M = 0.17, t(5) = 1.00, p > 0.05$, one-tailed), but was significant at 36 months of age ($M = 2.67, t(5) = 2.39, p < 0.05$, one-tailed).

In sum, our data show that both the English-speaking and Japanese-speaking children and mothers tended to cast given arguments over new ones in A position (subject role of transitive verbs). This provides support for Du Bois's (1987) constraint that the use of new A's is crosslinguistically avoided. Our data show slight differences between the English-speaking and Japanese-speaking mothers as to the location of new arguments. The English-speaking mothers tended to cast these in O position, whereas the Japanese-speaking mothers tended to cast them in S position. The English-speaking children

showed patterns similar to their input, and interestingly, the Japanese-speaking children showed patterns similar to the English speakers. This is further discussed below.

Distribution of New Lexical Arguments

Our analyses in the previous section clearly indicated that English-speaking and Japanese-speaking children and mothers avoided the use of new arguments, as well as lexical arguments, in A role position. In this section, we test our hypothesis that arguments that are simultaneously both new *and* lexical would be avoided in A role position, across age periods. Single-mean *t*-tests (one-tailed) were run to see whether the occurrence of new lexical arguments in each of the three grammatical roles differed significantly from zero and Tukey HSD tests (two-tailed) were run to see whether there were any changes across the two age periods. Mean frequencies and standard deviations of the children's and mothers' use of new lexical arguments in A, O, and S positions are shown in Table 10.

The English-speaking mothers never used new lexical arguments in A position at 21 months. Although the mean frequency of new lexical A's at 36 months ($M = 0.50$) differed significantly from zero ($t(5) = 2.24, p < 0.05$), new lexical arguments rarely occurred in A position, supporting our hypothesis. The mean frequency of new lexical S's did not differ significantly from zero at 21 months ($M = 0.33$, $t(5) = 1.58, p > 0.05$), but differed significantly at 36 months ($M = 1.00$, $t(5) = 2.74, p < 0.05$). Mean frequencies of new lexical A's did not differ significantly between 21 and 36 months ($Q(6, 25) = -3.15, p > 0.05$), indicating no change from one age period to the next. Nor did the mean frequencies of new lexical S's differ significantly between 21 and 36

months ($Q(6, 25) = -1.90, p > 0.05$). The mean frequencies of new lexical O's were significantly different from zero at both 21 ($M = 8.17, t(5) = 3.94, p < 0.05$) and 36 months of age ($M = 10.83, t(5) = 4.13, p < 0.05$).

The English-speaking children never ($M = 0.00$) used new lexical arguments in A position at either age period, consistent with our hypothesis. The mean frequencies of new lexical arguments in S position were not significantly different from zero at either age period (21 months: $M = 0.17, t(5) = 1.00, p > 0.05$; 36 months: $M = 0.33, t(5) = 1.00, p > 0.05$). New lexical O's were not significantly different from zero at 21 months ($M = 1.00, t(5) = 1.23, p > 0.05$), but were significantly different at 36 months of age ($M = 8.33, t(5) = 3.90, p < 0.05$). The mean frequency of new lexical O's differed significantly between 21 and 36 months of age ($Q(6, 25) = -5.52, p < 0.05$), indicating a change over time.

TABLE 10 ABOUT HERE

Consistent with our hypothesis, the Japanese-speaking mothers' use of new lexical arguments in A position did not differ significantly from zero at either age period (21 months: $M = 0.33, t(5) = 1.00, p > 0.05$; 36 months: $M = 0.17, t(5) = 1.00, p > 0.05$). By contrast, new lexical arguments occurring in S and O positions were significantly different from zero at both 21 months (new lexical S's: $M = 4.67, t(5) = 2.70, p < 0.05$; new lexical O's: $M = 1.00, t(5) = 3.87, p < 0.05$) and 36 months (new lexical S's: $M = 4.67, t(5) = 8.37, p < 0.05$; new lexical O's: $M = 1.33, t(5) = 2.70, p < 0.05$). New lexical

S's and O's did not differ significantly from each other at 21 months ($Q(6, 25) = -2.87, p > 0.05$), but did differ significantly at 36 months ($Q(6, 25) = -14.10, p < 0.05$).

The Japanese-speaking children never ($M = 0.00$) used new lexical arguments in A role at either age period, supporting our hypothesis regarding the avoidance of new lexical A's. The children used new lexical arguments in S position at both 21 ($M = 0.17$, $t(5) = 1.00, p > 0.05$) and 36 months of age ($M = 0.83$, $t(5) = 1.54, p > 0.05$), the mean frequencies of which were not significantly different from zero. New lexical O's only occurred at 36 months of age, and the mean frequency was significantly different from zero ($M = 2.17$, $t(5) = 4.54, p < 0.05$).

In sum, these results indicate that both the English-speaking and Japanese-speaking children and mothers tended to avoid casting new lexical arguments in A position, consistent with our Non-Lexical and Given A hypothesis. However, whereas the English-speaking children and mothers as well as the Japanese-speaking children reserved new lexical arguments for O position (object position of transitive verbs), the Japanese-speaking mothers tended to cast new lexical arguments in S position (subjects of intransitive verbs).

Language-Specific Referential Options

To test our hypothesis that non-lexical (null and pronominal) arguments in reference to given information are used in accordance with language-specific grammatical principles by 36 months of age, we performed a 2 (given null, given pronominal) x 2 (21 months, 36 months) repeated-measures ANOVA on frequency data for each of the four groups. We expected the English-speaking children to use

pronominal forms over null forms, but the Japanese-speaking children to continue using null forms, rather than pronominal ones. The mothers, by contrast, were not expected to modify their patterns from one age period to the next. Mean frequencies and standard deviations for the children's and mothers' use of null and pronominal arguments in reference to given information are shown in Table 11.

The two-way interaction between argument form and age for the English-speaking mothers was not significant ($F(1, 5) = 3.74, p > 0.05$), indicating that the mothers' patterns of use did not change between age periods. Consistent with an English-specific discourse-pragmatic strategy, the mothers used more pronominal forms than null forms in reference to given information across age periods, and this main effect was significant ($F(1, 5) = 78.17, p < 0.05$). Consistent with our hypothesis, a significant two-way interaction between argument form and age was found for the English-speaking children ($F(1, 5) = 17.54, p < 0.05$). At 21 months of age, pronominal forms ($M = 9.00$) were used almost as frequently as null forms ($M = 9.17$) in reference to given information, but by 36 months of age pronominal forms ($M = 38.33$) were used considerably more frequently than null forms ($M = 5.00$).

TABLE 11 ABOUT HERE

The two-way interaction between argument form and age was not significant for the Japanese-speaking mothers ($F(1, 5) = 0.12, p > 0.05$), indicating that the Japanese-speaking mothers did not change their patterns of use between age periods. A significant main effect of argument form was obtained ($F(1, 5) = 316.16, p < 0.05$). Consistent with

a Japanese-specific discourse-pragmatic strategy, the Japanese-speaking mothers used more null forms than pronominal forms in reference to given information across age periods. A significant two-way interaction between argument form and age was obtained for the Japanese-speaking children ($F(1, 5) = 12.48, p < 0.05$). The Japanese-speaking children used only a few null forms ($M = 2.00$) in reference to given information and never used pronominal forms ($M = 0.00$) at 21 months of age. However, consistent with our hypothesis, they used null forms most of the time at 36 months of age (null: $M = 27.50$; pronominal: $M = 1.67$).

In sum, the findings show that both groups of children conformed to language-specific grammatical principles by 36 months of age. Whereas the English-speaking children used pronominal arguments in reference to given information, the Japanese-speaking children continued to use null arguments. These language-specific grammatical constraints were provided in their respective inputs from as early as 21 months of age.

Discussion

Our goal in this paper was twofold: First, we sought to understand whether the discourse-pragmatic constraints underlying Du Bois's (1987) Preferred Argument Structure would be upheld developmentally and crosslinguistically. Previous studies looking at a discourse-pragmatic explanation of children's argument choices have shown promise in this respect, but none have empirically tested this. We thus conducted a study of PAS in English and Japanese, one overt and one null argument language, and analysed data at two time points in the children's development, 21 and 36 months of age. Secondly, we investigated the role played by parental input in children's learning of the

distributional patterns of arguments in syntax and discourse. We sought to understand how children come to learn language-universal discourse-pragmatic strategies, such as PAS constraints, as well as language-specific referential options, such as whether a language grammaticizes given information as null or pronominal.

Before we interpret our results with respect to the questions we ask in this paper, we first review the findings concerning the four PAS constraints. The quantity constraints as proposed by Du Bois (1987) were supported by our data: Both English-speaking and Japanese-speaking children and mothers tended to avoid using more than one new argument or more than one lexical argument per transitive clause. These findings support the One New Argument and One Lexical Argument constraints and indicate that the production of more than one new argument or more than one lexical argument per clause is extremely rare – even with children learning an overt argument language such as English, which had not been observed before. Likewise, Du Bois's syntactic role constraints were also supported by our data. Although language-specific variations in regards to overall frequency were observed for English and Japanese, the patterns are generally similar: Both English-speaking and Japanese-speaking children and mothers showed a general tendency to avoid casting new or lexical arguments in A position (subject position of transitive verbs), thus supporting the Non-Lexical A and Given A constraints. In addition, the analysis of arguments that are simultaneously both new and lexical provides more direct support for the two role constraints, furthermore indicating that casting new lexical arguments as transitive subjects is extremely rare crosslinguistically. Altogether, our PAS findings are consistent with those reported in the

literature for both child (e.g., Allen & Schröder, 2003; Clancy, 2003) and adult speakers (e.g., Du Bois, 1987).

However, a few inconsistencies with previous research were also observed, in particular with regards to the placement of lexical arguments. Although the Non-Lexical A Constraint was supported, the English-speaking mothers used lexical A's 10.5% of the time at 21 months (lexical A's decreased to 3.7% at 36 months). Similarly with the Japanese-speaking mothers, lexical A's occurred 23.2% of the time at 21 months and 24.3% at 36 months. By contrast, the adult Sakapultek speakers in Du Bois's (1987) study used lexical A's only 6.1% of the time and the Inuktitut-speaking children in Allen and Schröder's (2003) study produced lexical A's a mere 1.1% of the time. On the other hand, Clancy (2003) reported 12%–14% lexical A's by Korean-speaking children. Preliminary analyses indicate that the lexical A's produced by the English-speaking and Japanese-speaking mothers in the present study reflect the use of proper names or kinship terms in self- or addressee-reference (e.g., *Mommy will help you with that*, where a kinship term is used in place of a first-person pronoun). Almost half (46%) of the English-speaking mothers' and over 90% of the Japanese-speaking mothers' lexical A's were used in self- or addressee-reference. The use of proper names or kinship terms in A position is likely what affected the Non-Lexical A Constraint for the English-speaking mothers at 21 months and the Japanese-speaking mothers at 36 months, which were found to be significantly different from zero. Proper names or kinship terms might also have contributed to Clancy's elevated proportions of lexical A's. We are currently analysing the children's use of proper names and kinship terms and their relationship to PAS.

Also inconsistent with previous research were the English-speaking mothers' and children's use of new A's at 36 months: The mothers used new A's 9.1% of the time; the children 5.4% of the time. Although these turned out to be significant, the Given A Constraint was supported overall. By contrast, the Japanese-speaking mothers used new A's only 2% and 2.3% at 21 and 36 months, respectively. Du Bois (1987) reported 3.2% new A's by the Sakapultek speakers and Allen and Schröder (2003) reported 0.7% for the Inuktitut-speaking children; whereas Clancy (2003) reported 2%–4% for the Korean-speaking children. Preliminary analyses indicate that our English-speaking mothers' use of new A's reflect references to a third person, usually the father, who was not present during the interaction, and thus coded as "new" based on our coding criteria. We suspect that the English-speaking children's use of new A's similarly reflects references to a third person, and we are currently investigating this hypothesis.

Finally, our results showed an English-specific pattern for the three grammatical roles that seems to differ from previous studies. Our English data indicate that the A and S argument roles patterned alike with respect to the Non-Lexical A and Given A Constraints. In addition to avoiding the use of lexical or new A's, the English-speaking mothers and children similarly tended to avoid lexical or new S's. By contrast, Du Bois (1987) reported that S and O tended to pattern alike for adult Sakapultek speakers, in particular with respect to the representation of new information. A similar finding is reported by Allen and Schröder (2003) with respect to child Inuktitut. Our Japanese-speaking mothers' data also suggest the patterning of S with O for new information. The patterning of S with O (versus A) constitutes what Du Bois claims to be an ergative pattern of discourse and suggests it to be type-independent. That is, an ergative pattern is

predicted to underlie accusative languages (like English) as well, but an ergative pattern was not obtained for our English data; rather S patterned with A. Others have similarly reported the alignment of S with A for adult English narrative and conversational discourse (e.g., Kärkkäinen, 1996; Kumpf, 1992). Kärkkäinen (1996) suggests that the alignment of S with A in English is likely due to the grammaticization of English subjects as "light starting points," as proposed by Chafe (1994; 1996), in which case not only A, but S as well, function to accommodate given and non-lexical referents. Moreover, unlike Sakapultek, which is an ergative/absolutive language, English is nominative/accusative. Although Du Bois is likely correct in assuming a universal Preferred Argument Structure patterning, a universal ergative discourse patterning requires further study.

We now turn to the question we asked at the outset: How do children learn language-universal as well as language-specific discourse-pragmatic strategies, such as PAS? The findings reviewed above indicate that the English-speaking children showed sensitivity to all four PAS constraints from as early as, if not prior to, 21 months of age. Although the data for the Japanese-speaking children are inconclusive at 21 months, because of the overall lack of data, the children never showed violations of the constraints and, moreover, showed clear sensitivity to all four constraints by 36 months of age. These findings suggest that PAS strategies might be available to children from very early on in development. Secondly, both groups of children showed similar PAS patterns, indicating that the same discourse-pragmatic strategies are used regardless of the language the children are exposed to, whether null or overt argument. Given the crosslinguistic similarity and early sensitivity to PAS patterns, these findings suggest that discourse-pragmatic principles might be universal, as suggested by Du Bois (1987). It

should be noted, however, that even though neither the English-speaking nor the Japanese-speaking children showed violations of PAS constraints at early ages, the possibility remains that the reason no violations occurred is because the children simply omitted all (subject and object) arguments, resulting in the production of utterances that seem to suggest support of PAS strategies.

The children's data also show language-specific differences and these differences can only be accounted for by learning via parental input. First, language-specific referential options, such as whether to omit or pronominalize a given argument, were learned from exposure to patterns provided in parental input (see also Guerriero, Oshima-Takane, & Kuriyama, 2004). For instance, the English-speaking children, whose mothers used pronominal forms in reference to given information consistently between 21 and 36 months, learned to use pronominal forms by 36 months of age. At 21 months of age, the children had not shown differential use of null and pronominal arguments when making given references. By contrast, the Japanese-speaking children, whose mothers used null forms in reference to given information, also consistently between 21 and 36 months, learned that null forms were an acceptable means of reference by 36 months of age. Thus, by 36 months of age both English-speaking and Japanese-speaking children showed patterns strikingly similar to those of their mothers, indicating the influence of parental input.

Further support for the role of input can be seen in how each language grammaticizes lexical arguments. In avoiding the use of lexical arguments as transitive subjects, the English-speaking mothers tended to cast them as transitive objects. A similar pattern was observed with the English-speaking children. The Japanese-speaking

mothers, on the other hand, tended to show a more balanced distribution of lexical arguments across the three syntactic roles (though these were always less frequently cast as transitive subjects), and a similar pattern was observed with the Japanese-speaking children. Again, these language-specific patterns were not evident in the children's language when they were 21 months of age (although the findings are inconclusive with the Japanese-speaking children at 21 months of age), indicating that they learned the language-specific placement of lexical arguments via exposure to such patterns in their respective parental inputs.

The situation, however, was slightly different with respect to the location of new arguments. In avoiding the use of new arguments as transitive subjects, the English-speaking mothers tended to cast these as transitive objects. This pattern was similarly observed with the English-speaking children; and interestingly, with the Japanese-speaking children as well, but not with the Japanese-speaking mothers. The Japanese-speaking mothers, by contrast, tended to cast new arguments more frequently as intransitive subjects. This finding suggests a discrepancy between the Japanese-speaking children and their input, suggesting that input might not be as significant as we suggest. This discrepancy needs to be addressed, but first we address the inconsistent use of new arguments between the English-speaking and Japanese-speaking mothers.

One possibility that might account for this difference between the English-speaking and Japanese-speaking mothers is that surface syntactic differences could have resulted from the genre of data under analysis here. For instance, Du Bois (1987) reported on adult narrative data, whereas ours is spontaneous speech. But since our Japanese data are also spontaneous speech, the differences we observed with respect to the English-

speaking and Japanese-speaking mothers' casting of new information are likely not due to the genre of the data.

Rather, differences are likely due to how each language grammaticizes new information. Du Bois (1987) reported that adult Sakapultek speakers similarly tended to cast new arguments as intransitive subjects almost as often as transitive objects (intransitive subjects contained 22.5% new information; transitive objects contained 24.7% new information),[18] and suggested that the high incidence of new intransitive subjects stemmed from the discourse function of intransitive verbs in managing information flow. According to Du Bois, due to the restriction imposed by transitive subjects in allowing only given arguments (the Given A Constraint), the intransitive subject role is frequently selected to introduce new information, in particular when the new referent is a human protagonist. Animate referents typically appear as transitive subjects (whereas inanimate ones appear as transitive objects). This is referred to as the "animacy hierarchy" and is a discourse feature that has been observed in several languages (e.g., Allen & Schröder, 2003; Clancy, 2003; Du Bois, 1987). Thus, in order to introduce a new (animate) human referent without violating the Given A Constraint, the intransitive frame is often preferred. These intransitive frames are often copular in

[18] Note that Du Bois (1987) differentiated between referents mentioned for the first time ("new") versus those mentioned more than 20 utterances previously ("accessible"). Accessible referents are considered to be "semi-active," in the sense that they have been retrieved from long-term memory, but are not the current focus of attention (see Chafe, 1987; 1994; 1996). The figures reported above are in regards to new referents only; however, combining together new and accessible referents, the Sakapultek speakers used new arguments as intransitive subjects 27.5% of the time and as transitive objects 34.7% of the time. In our research, we follow Allen and Schröder (2003) and Clancy (2003), who combined accessible and new referents. Reasons for this are, first, because the number of accessible referents overall is quite small, and second, it is unclear whether referents mentioned more than 20 utterances previously are retained in children's memories as they would be for adults.

structure (e.g., *Here is a boy*), but may also contain a neutral verb such as "come," "arrive," or "appear" (Du Bois, 1987, p. 831). Once the new referent has been introduced, the referent, now given, can be re-introduced as a transitive subject. In other words, casting new information as intransitive subjects, which do not have discourse restrictions, is due to the compatibility of intransitive verbs with the PAS constraints underlying the flow of information in discourse. We are currently investigating the relationship between animacy and the distribution of sentence arguments in discourse.

Clancy (2003) conducted qualitative analyses on four of the most frequent intransitive verbs in her child Korean data set, which accounted for 56%–67% of new arguments, and reported that the high rate of new (and lexical) arguments was due to the function of intransitives for referent introductions (consistent with Du Bois's suggestion). The verbs served interactive functions: They focussed interlocutors' attention to new persons, pictures, or objects for the child and caregiver to interact with, talk about, or act upon. Moreover, Clancy's finding indicates that the discourse function of intransitive verbs, as suggested by Du Bois, is also characteristic of young (Korean-speaking) children's speech.

As already alluded to, it is not unexpected that English and Japanese show language-specific differences with respect to the grammaticization of discourse information; since the two languages are typologically different. Whereas Japanese speakers cast new information in intransitive subject role, English speakers cast it in transitive object role. Kärkkäinen (1996), who investigated PAS strategies in adult American English conversational discourse, reported that new information tended to occur predominantly as transitive objects – similar to the pattern observed with the

English speakers in the present study. Likewise, Kumpf (1992) reported that new arguments appeared predominantly as transitive objects in adult English conversational discourse. Like their mothers, the English-speaking children in the present study tended to cast new information as transitive objects. The interesting question, however, is why the Japanese-speaking children did not show a pattern similar to that of their mothers, whereby new information was cast in intransitive subject role.

Inconsistent with their mothers, the Japanese-speaking children tended to cast new arguments in transitive object position. How did the children learn this pattern if it was not provided in their input? In addition, the Japanese-speaking mothers and children showed discrepancies in terms of the proportions of transitive and intransitive frames. The Japanese-speaking mothers used intransitive verbs 71.7% at 21 months and 63.5% at 36 months (see Table 3), whereas the children used a more balanced distribution of transitive to intransitive verbs (51.6% intransitive and 48.4% transitive verbs at 36 months of age). Thus, how did the children learn to use transitive frames without much exposure to them? We suggest that the mothers did provide the children with this type of input, but it was not included in our analyses. Recall that our stringent exclusionary criteria only allowed sentences with nominal referring expressions. Thus, transitive verbs taking sentential complements or activity-like objects were excluded in their entirety – both subject and object arguments. Furthermore, we excluded all imperative utterances from analysis. This was so because of the null subject argument that is contained within an imperative sentence, and we did not want to over-inflate the number of null subjects in our data. The majority of exclusions, in fact, consisted of imperative utterances. However, upon a re-examination of our data, it was revealed that the Japanese-speaking

mothers did, in fact, show a more balanced distribution of transitive and intransitive clauses. If proportions are tabulated based on all coded data, without exclusions, the mothers show a rate of 55.1% and 52.6% intransitive clauses at 21 and 36 months, respectively. These figures are more in line with the figures reported by Clancy (2003) for the Korean-speaking children. Likewise, Du Bois (1987) reported that the adult Sakapultek speakers showed only slightly more intransitive clauses (60%) over transitive ones (40%).

Likewise, without exclusions, the Japanese-speaking children used intransitive clauses 43.6% of the time at 36 months of age, a figure more similar to that of the Japanese-speaking mothers, and also to that reported in the literature (e.g., Clancy, 2003; Du Bois, 1987). By excluding some of our data, the results had shown a misleading proportion of transitive to intransitive clauses and, more importantly, discrepant patterns between the Japanese-speaking children and their mothers. It thus turns out that the Japanese-speaking children are as similar to the Japanese-speaking mothers as the English-speaking children are to the English-speaking mothers, indicating that language-specific variations are learned via patterns observed from parental input. These findings also indicate that imperative speech needs further investigation. Given the frequent use of imperatives in adult speech to children, it is likely that imperative structures provide a good source of input regarding the morphological representation and syntactic distribution of sentence arguments. We are currently investigating this possibility.

In sum, our study has shown that children learning an overt argument language, such as English, show similar PAS patterns to children learning null argument languages. Previous research looking at childhood PAS has examined null argument languages, such

as Inuktitut and Korean; thus, our present investigation with English child language contributes new data to this line of research. This is the first systematic and crosslinguistic investigation to employ the same methodology to show that discourse-pragmatic constraints generalize across both null and overt argument languages. Secondly, we have shown that PAS sensitivity seems to be in place from early stages of language development, although we are currently investigating a second interpretation of the data that early PAS sensitivity might be due to the pervasive omission of all arguments at early stages. Finally, this study has shown that argument representation and its distribution in surface syntax comes about from the interaction of language-universal principles as well as language-specific linguistic strategies – strategies which are learned from input and which develop over time.

Table 1
Dimensions and constraints of Preferred Argument Structure

	Grammar	Pragmatics
Quantity	One Lexical Argument Constraint: Avoid more than one lexical argument per clause.	One New Argument Constraint: Avoid more than one new argument per clause.
Role	Non-Lexical A Constraint: Avoid lexical A's.	Given A Constraint: Avoid new A's.

Note. Adapted from Du Bois (1987). Du Bois used the following notation: A = transitive subject; O = transitive object; S = intransitive subject.

Table 2
The Structures of Japanese and English

	Japanese	English
Word order	SOV	SVO
Argument omission	optional	not permitted
Verbal agreement	none	some
Case	case markers	marked by word order

Table 3
Frequencies of Verbs and Arguments Analysed

	Children		Mothers	
	21 months	36 months	21 months	36 months
English speakers				
Total Verbs	84	171	295	223
Transitive	58 (69.0%)	148 (86.5%)	201 (68.1%)	150 (67.3%)
Intransitive	26 (31.0%)	23 (13.5%)	94 (31.9%)	73 (32.7%)
Total Arguments	142	319	496	373
Japanese speakers				
Total Verbs	16	155	226	274
Transitive	1 (6.3%)	75 (48.4%)	64 (28.3%)	100 (36.5%)
Intransitive	15 (93.7%)	80 (51.6%)	162 (71.7%)	174 (63.5%)
Total Arguments	17	230	290	374

Note. $N = 6$ speakers in each age group. Proportions of transitive and intransitive clauses appear in parentheses. Transitive verbs contain two arguments each; hence, the total number of arguments is based on the following formula: (number of transitive verbs x 2) + number of intransitive verbs. Only transitive verbs with two coded arguments were included for analysis.

Table 4
Transitivity and Number of New Arguments per Clause: Mean Percentages and Standard Deviations

Number of Arguments		Transitive Clauses			Intransitive Clauses	
		Zero	One	Two	Zero	One
English Speakers: Mothers						
21m	Mean	22.17 (67%)	11.17 (32%)	0.17 (1%)	14.50 (92%)	1.17 (8%)
	SD	7.55 (8%)	5.81 (8%)	0.41 (2%)	4.89 (10%)	1.33 (10%)
36m	Mean	11.50 (48%)	12.00 (46%)	1.50 (5%)	9.83 (82%)	2.33 (18%)
	SD	3.02 (16%)	5.80 (16%)	1.87 (6%)	4.71 (14%)	1.63 (14%)
Children						
21m	Mean	7.60 (59%)	4.00 (41%)	0.00 (0%)	3.67 (79%)	0.67 (21%)
	SD	9.76 (13%)	2.65 (13%)	0.00 (0%)	2.50 (33%)	1.21 (33%)
36m	Mean	10.33 (41%)	14.17 (58%)	0.17 (2%)	3.80 (81%)	0.80 (19%)
	SD	8.19 (17%)	10.57 (19%)	0.41 (4%)	3.70 (26%)	1.30 (26%)
Japanese Speakers: Mothers						
21m	Mean	7.83 (69%)	2.67 (31%)	0.17 (1%)	18.50 (64%)	8.50 (36%)
	SD	7.00 (17%)	2.94 (17%)	0.41 (2%)	12.37 (14%)	4.04 (14%)
36m	Mean	14.50 (86%)	2.00 (13%)	0.17 (1%)	21.17 (72%)	7.83 (28%)
	SD	7.82 (10%)	0.89 (7%)	0.41 (3%)	5.71 (9%)	1.83 (9%)
Children						
21m	Mean				4.67 (97%)	0.33 (3%)
	SD	--	--	--	6.35 (4%)	0.58 (4%)
36m	Mean	7.67 (54%)	4.83 (46%)	0.00 (0%)	10.67 (83%)	2.67 (17%)
	SD	5.43 (20%)	1.72 (20%)	0.00 (0%)	4.76 (14%)	2.73 (14%)

Note. $N = 6$ speakers in each age group, unless otherwise indicated. Mean frequencies and standard deviations appear within parentheses. $N = 5$ for English-speaking children's transitive verbs at 21 months and intransitive verbs at 36 months. $N = 3$ for Japanese-speaking children's intransitive verbs at 21 months. Data for transitive clauses are not given for Japanese-speaking children at 21 months because only one child produced one transitive verb at this age.

Table 5
Transitivity and Number of Lexical Arguments per Clause: Mean Percentages and Standard Deviations

Number of Arguments		Transitive Clauses			Intransitive Clauses	
		Zero	One	Two	Zero	One
English Speakers: Mothers						
21m	Mean	13.33 (39%)	18.33 (54%)	1.83 (7%)	14.17 (90%)	1.50 (10%)
	SD	7.00 (11%)	8.02 (11%)	1.72 (8%)	4.71 (12%)	1.64 (12%)
36m	Mean	9.17 (39%)	14.83 (58%)	1.00 (4%)	10.67 (89%)	1.50 (12%)
	SD	2.99 (17%)	6.27 (19%)	1.55 (5%)	4.68 (10%)	1.38 (10%)
Children						
21m	Mean	7.80 (84%)	3.60 (15%)	0.20 (1%)	4.00 (89%)	0.33 (11%)
	SD	4.97 (21%)	6.95 (20%)	0.45 (1%)	2.00 (20%)	0.52 (20%)
36m	Mean	13.33 (50%)	11.17 (49%)	0.17 (1%)	4.20 (94%)	0.40 (6%)
	SD	9.99 (12%)	6.55 (12%)	0.41 (1%)	3.56 (13%)	0.89 (13%)
Japanese Speakers: Mothers						
21m	Mean	5.33 (48%)	4.83 (42%)	0.50 (10%)	17.50 (69%)	9.50 (31%)
	SD	4.97 (5%)	3.76 (22%)	0.84 (20%)	8.46 (11%)	6.95 (11%)
36m	Mean	8.50 (50%)	6.83 (41%)	1.33 (9%)	19.17 (65%)	9.83 (35%)
	SD	5.24 (15%)	3.37 (7%)	1.75 (14%)	6.01 (12%)	3.31 (12%)
Children						
21m	Mean				3.00 (85%)	2.00 (15%)
	SD	--	--	--	3.46 (27%)	3.46 (27%)
36m	Mean	6.67 (47%)	5.17 (48%)	0.67 (5%)	10.67 (83%)	2.67 (17%)
	SD	5.32 (19%)	2.86 (22%)	0.82 (6%)	4.46 (16%)	2.88 (16%)

Note. $N = 6$ speakers in each age group, unless otherwise indicated. Mean frequencies and standard deviations appear within parentheses. $N = 5$ for English-speaking children's transitive verbs at 21 months and intransitive verbs at 36 months. $N = 3$ for Japanese-speaking children's intransitive verbs at 21 months. Data for transitive clauses are not given for Japanese-speaking children at 21 months because only one child produced one transitive verb at this age.

Table 6
Distribution of Non-Lexical and Lexical Arguments across Grammatical Roles: English Speakers

	Mean Frequencies		Proportion Lexical
	Non-Lexical	Lexical	
Mothers: 21 months			
S	14.17 (4.71)	1.50 (1.64)	10.1% (12%)
A	30.67 (13.78)	2.83 (2.32)	10.5% (10%)
O	14.33 (7.17)	19.17 (7.73)	57.8% (12%)
36 months			
S	10.67 (4.68)	1.50 (1.38)	11.5% (10%)
A	24.00 (5.55)	1.00 (1.55)	3.7% (5%)
O	9.17 (2.99)	15.83 (6.62)	61.4% (17%)
Children: 21 months			
S	4.00 (2.00)	0.33 (0.52)	11.1% (20%)
A	9.33 (11.08)	0.33 (0.82)	1.2% (3%)
O	6.67 (5.82)	3.00 (6.39)	15.3% (20%)
36 months			
S	3.50 (3.62)	0.33 (0.82)	5.7% (13%)
A	24.50 (16.33)	0.17 (0.41)	0.4% (1%)
O	13.33 (9.99)	11.33 (6.71)	49.6% (12%)

Note. $N = 6$ speakers in each age group, unless otherwise indicated. Standard deviations appear within parentheses. $N = 5$ for the children's A and O arguments at 21 months and S arguments at 36 months.

Table 7
Distribution of Non-Lexical and Lexical Arguments across Grammatical Roles: Japanese Speakers

	Mean Frequencies		Proportion Lexical
	Non-Lexical	Lexical	
Mothers: 21 months			
S	17.50 (8.46)	9.50 (6.95)	31.4% (11%)
A	7.50 (5.39)	3.17 (4.54)	23.2% (26%)
O	8.00 (8.72)	2.67 (2.16)	38.4% (24%)
36 months			
S	19.17 (6.01)	9.83 (3.31)	34.6% (12%)
A	12.67 (6.80)	4.00 (3.10)	24.3% (17%)
O	11.17 (6.71)	5.50 (2.35)	34.4% (14%)
Children: 21 months			
S	1.50 (2.74)	1.00 (2.45)	15.4% (27%)
A	--	--	--
O	--	--	--
36 months			
S	10.67 (4.46)	2.67 (2.88)	16.5% (16%)
A	9.67 (5.13)	2.83 (1.94)	23.5% (11%)
O	8.83 (5.64)	3.67 (1.51)	34.3% (15%)

Note. $N = 6$ speakers in each age group, unless otherwise indicated. Standard deviations appear within parentheses. $N = 3$ for the children's S arguments at 21 months. Data are not given for the children's A and O arguments because only one child produced one transitive verb at this age.

Table 8
Distribution of Given and New Arguments across Grammatical Roles: English Speakers

	Mean Frequencies		Proportion New
	Given	New	
Mothers: 21 months			
S	14.50 (4.89)	1.17 (1.33)	8% (10%)
A	33.17 (12.58)	0.33 (0.52)	1% (2%)
O	22.33 (7.74)	11.17 (5.71)	32.3% (8%)
36 months			
S	9.83 (4.71)	2.33 (1.63)	18.4% (14%)
A	22.67 (5.47)	2.33 (2.66)	9.1% (9%)
O	12.33 (3.08)	12.67 (6.83)	48.2% (18%)
Children: 21 months			
S	3.67 (2.50)	0.67 (1.21)	20.8% (33%)
A	9.67 (11.86)	0.00 (0.00)	0% (0%)
O	6.33 (9.27)	3.33 (2.88)	41.2% (13%)
36 months			
S	3.17 (3.66)	0.67 (1.21)	18.6% (26%)
A	24.17 (17.01)	0.50 (0.55)	5.4% (7%)
O	10.67 (7.92)	14.00 (10.71)	55.4% (14%)

Note. $N = 6$ speakers in each age group, unless otherwise indicated. Standard deviations appear within parentheses. $N = 5$ for the children's A and O arguments at 21 months and S arguments at 36 months.

Table 9
Distribution of Given and New Arguments across Grammatical Roles: Japanese Speakers

	Mean Frequencies		Proportion New
	Given	New	
Mothers: 21 months			
S	18.50 (12.37)	8.50 (4.04)	36.4% (14%)
A	10.17 (7.99)	0.50 (1.23)	2% (5%)
O	8.17 (7.49)	2.50 (2.59)	29.9% (17%)
36 months			
S	21.17 (5.71)	7.83 (1.84)	27.8% (9%)
A	16.33 (7.94)	0.33 (0.52)	2.3% (4%)
O	14.67 (7.87)	2.00 (1.10)	13.5% (10%)
Children: 21 months			
S	2.33 (4.76)	0.17 (0.41)	2.6% (4%)
A	--	--	--
O	--	--	--
36 months			
S	10.67 (4.76)	2.67 (2.73)	17.1% (14%)
A	12.50 (5.96)	0.00 (0.00)	0% (0%)
O	7.67 (5.43)	4.83 (1.72)	45.9% (20%)

Note. $N = 6$ speakers in each age group, unless otherwise indicated. Standard deviations appear within parentheses. $N = 3$ for the children's S arguments at 21 months. Data are not given for the children's A and O arguments because only one child produced one transitive verb at this age.

Table 10
Mean Frequencies of New Lexical Arguments in Each Grammatical Role

	Mothers		Children	
Role	21 months	36 months	21 months	36 months
English speakers				
S	0.33 (0.52)	1.00 (0.89)	0.17 (0.41)	0.33 (0.82)
A	0.00 (0.00)	0.50 (0.55)	0.00 (0.00)	0.00 (0.00)
O	8.17 (5.08)	10.83 (6.43)	1.00 (2.00)	8.33 (5.24)
Japanese speakers				
S	4.67 (4.23)	4.67 (1.37)	0.17 (0.41)	0.83 (1.33)
A	0.33 (0.82)	0.17 (0.41)	0.00 (0.00)	0.00 (0.00)
O	1.00 (0.63)	1.33 (1.21)	0.00 (0.00)	2.17 (1.17)

Note. $N = 6$. Standard deviations appear within parentheses.

Table 11
Given Information: Mean Frequencies of Null and Pronominal Argument Forms

Argument Form	Mothers		Children	
	21 months	36 months	21 months	36 months
English speakers				
Null	10.83 (7.52)	3.83 (2.64)	9.17 (5.57)	5.00 (4.05)
Pronominal	70.83 (23.04)	60.00 (17.07)	9.00 (19.14)	38.33 (24.97)
Japanese speakers				
Null	41.17 (31.93)	48.33 (20.87)	2.00 (2.68)	27.50 (11.52)
Pronominal	1.83 (2.23)	3.33 (2.25)	0.00 (0.00)	1.67 (1.21)

Note. $N = 6$. Standard deviations appear within parentheses.

CHAPTER 5

General Discussion:

The Development of Argument Representation

The main objective underlying the three studies undertaken in this thesis was to investigate how children learn argument representation. Specifically, I sought to investigate how children learn to use the various referential expressions available in the target language, namely null, pronominal, and lexical expressions. The question, in essence, was two-fold: First, are children's early null arguments random or do they follow a systematic pattern, and second, given that children start off by omitting sentence arguments in early syntactic production, how do they come to learn language-specific conventions of argument representation, such as whether the target language is one which does or does not permit optional argument omission? That is, if the target language does not permit argument omission, a language such as English, how do children come to learn that pronominal forms are to be used instead; or if the target language does permit argument omission, such as Japanese, how do children come to learn that argument omission is a grammatically acceptable? Another objective of the thesis was to investigate whether a discourse-pragmatic approach could adequately account for children's patterns of argument omission and production across developmental periods, including the null argument period, since the competence- and performance-based approaches did not adequately support the observed data (refer to the General Introduction). Because the specific findings of each study are interpreted in the respective

discussion sections, this chapter is confined to a discussion of the findings of the three studies taken together as they relate to the main objectives of the thesis. First, a summary of the major findings is given.

The first and second studies (*The Development of Argument Representation in English and Japanese Child Language: A Discourse-Pragmatic Perspective* and *Children's Referential Choices: The Role of Non-Linguistic Pragmatic Information*) looked at the two-dimensional relationship between an argument's referential status (given versus new) and its morphological form (null, pronominal, or lexical). Based on a discourse-pragmatic approach, it was expected that given information would be omitted or pronominalized (depending on the grammatical conventions of the target language) and new information would be lexicalized. Findings from the first study revealed that the children, whether they were learning English or Japanese, did not seem to be sensitive to discourse-pragmatic principles early on in development (i.e., at 21 months of age) because the children tended to omit all arguments, whether given or new. In order to track developmental changes, the children were observed at two time points, 21 and 36 months of age, and by 36 months of age both groups showed expected discourse-pragmatic patterns for the referencing of given information: Given referents were pronominalized by the English-speaking children, but omitted by the Japanese-speaking children.

Because the findings from the first study also revealed extensive individual differences, the children's development was tracked by MLU, rather than age, in the second study. Moreover, in addition to the referential status and morphological form, the second study also investigated whether the referential expression was accompanied by

non-linguistic pragmatic information, such as pointing or touching gestures. Findings revealed that the English-speaking children showed mastery of English-specific referential conventions (i.e., the use of pronominal forms in reference to given information) between MLU 2.00 and 3.99 (between 24 and 31 months of age). Furthermore, if non-linguistic pragmatic correlates are taken into account, the children showed sensitivity from as early as MLU 1.00 (between 20 and 22 months of age) because they tended to supplement non-lexical references to new information, such as references made with a null or pronominal argument, with additional pragmatic information such as gestures; thus indicating sensitivity to discourse-pragmatic principles from quite early on. The Japanese-speaking children, on the other hand, continued to make references to new information that were inconsistent with a discourse-pragmatic strategy (i.e., the use of non-lexical forms) even by MLU 4.00 (between 33 and 36 months of age). Nor did they consistently supplement these non-specific forms with additional non-linguistic pragmatic information. These findings seem to indicate that the Japanese-speaking children were not sensitive to discourse-pragmatic principles of communication even by MLU 4.00.

The third study (*The Syntactic Distribution of Sentence Arguments in Child Language: The Role of Universal and Language-Specific Factors*) looked at the three-dimensional relationship between an argument's referential status (given versus new), its morphological form (null, pronominal, or lexical), and its syntactic location (transitive subject, transitive object, or intransitive subject). It was expected that the children's spoken utterances would conform to Du Bois's (1987) four Preferred Argument Structure (PAS) constraints. Findings revealed that both the English-speaking and Japanese-

speaking children tended to avoid casting new or lexical arguments as transitive subjects (supporting the Given A and Non-Lexical A Constraints) and tended to avoid using more than one new or lexical argument per (transitive) clause (supporting the One New Argument and One Lexical Argument Constraints). Consistent with previous studies of null argument child languages, the English-speaking children showed sensitivity to the four PAS constraints from very early in language development, at least as early as 21 months of age. Although the data with the Japanese-speaking children were found to be inconclusive at 21 months (the children produced very little at this age), the children showed clear sensitivity by 36 months of age.

Taken together, the findings from the three studies address several issues in children's learning of argument representation, namely, (1) the role played by language-universal discourse-pragmatic principles versus language-specific discourse-pragmatic strategies, (2) the role played by input, and (3) the adequacy of a discourse-pragmatic account in explaining children's referential choices. Each is discussed below.

Language-Universal and Language-Specific Discourse-Pragmatic Patterns

According to Du Bois (1987), Preferred Argument Structure constraints represent statistical tendencies underlying universal principles of discourse communication. These principles are presumed to motivate the distribution of arguments within syntax for speakers of all languages, null and overt argument alike. Secondly, under a discourse-pragmatic approach, the morphological realization of arguments is based on the informativeness status of discourse referents. Referents high in informativeness status

(i.e., new referents) tend to be lexicalized, whereas those low in informativeness status (i.e., given referents) tend to be non-lexicalized. If a language allows null arguments as a grammatical option, the referential expression typically results as an omitted argument; otherwise the argument is pronominalized. Thus, referential choice is presumed to be motivated by universal pragmatic principles of discourse; principles which are predicated upon the informativeness status of discourse referents.

Working concurrently with language-universal principles are the language-specific principles. Although the principles underlying referential choices are language-universal, the actual morphological realization of arguments is language-specific, in particular with respect to the realization of given information. The representation of a given referent as an omitted versus a pronominal argument is a feature based on the grammatical conventions of a specific language. A null argument language, such as Japanese, has both null and pronominal options available. In such a language, children have three referential options to learn, namely null, pronominal, and lexical forms. On the other hand, children learning English, an overt argument language, have only two options to learn, these being pronominal and lexical forms, since null arguments are not permitted.[19] Moreover, because English and Japanese are typologically different, the learning task differs between the two groups of children: The task for the Japanese-learning child involves learning that Japanese allows null arguments; whereas the task for

[19] Strictly speaking, even for children learning overt argument languages there are instances in which null arguments are permitted, either for grammatical or pragmatic reasons, such as when an imperative verb is used or when replying to a question by omitting the subject (e.g., "*What are you doing?*" "*Reading.*"). But these are specific instances, and as a rule, null arguments are not permitted in overt argument languages.

the English-learning child involves learning that argument omission is grammatically unacceptable, and as such, that pronominalization is to be used instead. Either way, the learning task involves mastering which discourse contexts call for each type of referential form.

Results from the third study provide support for the proposal that Preferred Argument Structure constraints represent universal tendencies, as it was found that the four PAS constraints were observed crosslinguistically, in both Japanese and English, and with both children and mothers alike. Although language-specific differences were obtained between the English-speaking and Japanese-speaking groups, violations of the constraints were never observed. This finding, that none of the four PAS constraints was ever violated, suggests that these principles might be available to children from very early in development, perhaps even before children produce their first words. However, it should be noted that the absence of violations might simply be an artifact of children's immature processing capabilities. Because very young children cannot produce more than one or two words at a time, the likely result is that the majority of clauses will not contain more than one (lexical or new) argument. In addition, the finding that the children avoided the use of (new or lexical) A's could simply be an artifact of generalized and widespread argument omission. In fact, the findings reported here show that the children generally tended to omit all arguments at early stages of linguistic development. These things considered, support for early sensitivity to PAS constraints is inconclusive at this point.

The second set of discourse-pragmatic principles, namely, the lexicalization of new information versus the non-lexicalization of given information, do not seem to be in

place early on in development. It had been hypothesized that children's argument choices would reflect sensitivity to the referential status of discourse referents early on in development, and, as such, it was predicted that null arguments would be in reference to given information only. However, the results revealed that neither the English-speaking nor the Japanese-speaking children distinguished between given and new information at early stages. The children tended to use null forms indiscriminately to refer to either type of information.

The English-speaking children eventually showed the predicted discourse-pragmatic patterns, whereby new information is lexicalized and given information is pronominalized. However, the Japanese-speaking children continued to make referential choices that seemed inconsistent with discourse-pragmatic principles throughout development, in particular with the referencing of new information. This crosslinguistic difference was unexpected and seems to be at odds with previous studies of null argument child languages, namely Korean and Inuktitut, which report that children show sensitivity to discourse-pragmatic principles when making referential choices (Allen, 1997; 2000; 2003; Clancy, 1993; 1997; Clancy, 2003). However, the Inuktitut-speaking children may have been older than the children in the present set of studies. Indeed, one of the Korean-speaking children who was less linguistically advanced tended to use null forms in reference to new information (also Skarabela & Allen, 2002). As will be more fully discussed in the next section, it is likely that the children in the studies reported here were simply following the model given in their input and, thus, that language-specific input patterns might override presumably language-universal ones.

The finding that the children did not discriminate between given and new discourse referents early on in development suggests that there is an element of learning involved in choosing which aspects of the discourse situation to verbally encode and which to leave presupposed. It seems that the learning required on the part of the child extends beyond simply learning the grammatical conventions for the referencing of given information. The results obtained here suggest that children must learn how to make appropriate reference to items of varying degrees of informativeness. In other words, children need to learn when it is appropriate to use each referential form, or which discourse situations necessitate a lexical mention versus which necessitate a non-lexical mention.

The finding that children do not come equipped knowing the referential distinction between given and new information seems to challenge Greenfield's (1979; also Greenfield & Smith, 1976) work showing that children at the one-word stage use the principle of informativeness in choosing which situational element to verbally produce and which to leave unexpressed. However, the principle of informativeness in early child language, as Greenfield suggested it, is based on the perception of uncertainty in the discourse situation from the child speaker's point of view. An element that is "new" from an adult's perspective might not necessarily be informative from a child's point of view. The findings obtained indicate that the English-speaking and Japanese-speaking children had yet to learn which aspects of the discourse situation were the most informative, thus necessitating overt mention, from the listener's point of view.

In sum, the picture which emerges from these studies is that argument representation comes about from the interrelationship between language-universal

principles and language-specific strategies. Language-universal principles, such as discourse-pragmatic strategies, combine with language-specific grammatical conventions in the selection of a morphological form and its syntactic distribution. Language-universal constraints are crosslinguistic, although they are amenable to incorporating language-specific tendencies. It is unclear at this point whether children's early conformity to PAS syntactic structure results from actual sensitivity to discourse-pragmatic constraints or whether the resulting output is merely an artifact of widespread argument omission. However, it is clear that children do not know the referential distinction between given and new information from the start of language production. On the other hand, language-specific referential options and strategies must be learned and developed gradually over time, and clearly, this is where the role of input comes into play.

Input

Analyses of the children's input indicate a relationship between the children's referential strategies and those of their mothers. The results from the first study showed that the Japanese-speaking children continued to use non-lexical forms (mainly null arguments) in reference to new information at 36 months of age. Examination of the Japanese-speaking mothers' data revealed that the mothers used this pattern when their children were 21 months of age, which suggests that the children may have learned this referential strategy from their input. The second study, which analysed the use of non-linguistic pragmatic information, revealed that the Japanese-speaking children's particular use (and non-use) of gestures could be accounted for by the patterns provided

in their respective inputs. Other researchers, who have reported similar use of null arguments in reference to new information, have shown that the speakers are typically engaged in joint attention, thus obviating the need for a lexical mention (Skarabela & Allen, 2002). Finally, the third study showed that even while abiding within PAS constraints, the Japanese-speaking children tended to distribute lexical arguments evenly across the three grammatical roles of transitive subject, transitive object, and intransitive subject; a pattern that was similarly observed with the Japanese-speaking mothers. Altogether, the findings indicate that the Japanese-speaking children's referential strategies were patterned after those of the Japanese-speaking mothers.

The influence of parental input was also observed with the group of English-speaking children, whose referential choices in fact conformed to predicted discourse-pragmatic principles throughout. There is no doubt that the children learned to pronominalize given information by observing the use of pronouns in the language of the English-speaking mothers. Because the choice between pronominalization and omission is a language-specific grammatical convention, it is obvious that this feature is learned from parental input. Findings obtained in the second study (i.e., Chapter 3) provide evidence that children master pronominalization of given information between MLU 2.00 and 3.99. However, more significant evidence of the influence of input was observed in the children's use of non-linguistic pragmatic correlates. The English-speaking children tended to accompany non-lexical (null or pronominal) references to new information with additional pragmatic information such as points and touches, and their mothers showed an identical strategy. The third study showed that, in avoiding the casting of lexical arguments in transitive subject position, the English-speaking children tended to cast

these predominantly as transitive objects; a pattern that was strikingly similar to the pattern used by their mothers.

The picture emerging from this data is that even though speakers might be constrained by universal discourse-pragmatic principles, there is still room for crosslinguistic differences to occur and that such crosslinguistic differences develop due to differences observed from environmental input. For instance, the casting of lexical arguments points to a language-specific crosslinguistic difference. These findings indicate that children perceive – and thus learn – language-specific syntactic structure by paying attention to patterns regularly provided in their input. This proposal, that children learn argument representation via input, is consistent with a growing body of literature indicating that input plays a more significant role in language acquisition than is typically assumed (e.g., by generative linguists and others proposing innate structures specifically dedicated to language learning). Studies investigating the acquisition of word order which have shown that children learn verbs on an item-based manner (what Tomasello (1992a) refers to as the "verb-island hypothesis") similarly attribute a significant role to the influence of input. These studies have shown that children's grammatical knowledge is initially structured around specific lexical items and that these first verbs are typically those that children hear most frequently in their input. Only later, by paying attention to recurring syntactic patterns, children construct generalizations of more abstract syntactic structures (Akhtar, 1999; Cameron-Faulkner, Lieven, & Tomasello, 2003; Lieven, Pine, & Baldwin, 1997; Theakston, Lieven, Pine, & Rowland, 2004; Tomasello, 1992a; 2000a; 2000b). For example, learning the abstract structure for transitive (SVO) verbs in English is only possible after children have learned many different verbs used in the transitive

frame, as well as hearing those same verbs used in other contrasting frames (e.g., the intransitive SV frame). However, until children have formed an abstract syntactic understanding of transitive verbs, they will tend to replicate the transitive frames modelled with specific verbs heard in their input.

In sum, the studies reported in this thesis point to the particularly influential nature of parental input to the learning of argument representation. The results indicate that both groups of children learned the language-specific grammatical requirements of their target languages via patterns observed from parental input. In particular for the Japanese-speaking children, who showed what seemed to be unconventional discourse-pragmatic strategies, the same patterns were observed in the children's input. The patterns to which children pay attention include both linguistic and pragmatic ones. This is because in order to learn the referential distinction between given and new information, children need to perceive the various discourse and situational contexts to which lexical and non-lexical forms are being applied. Thus, in combination with the discourse-pragmatic context, children use the corresponding syntactic patterns heard in their input to learn not only argument representation, but argument structure as well.

The learning of argument representation via experience with linguistic and pragmatic input implies that young children have access to powerful learning mechanisms with which to perform distributional analyses of their environmental input (Akhtar, 1999; Goldberg, 1998; Hoff-Ginsberg, 1985). Studies show that young language-learners are adept at detecting patterns in their input (Naigles, 2002). It has been shown that infants can detect statistical regularities in recurring sound sequences after a mere two-minute listening exposure to the sounds from as early as 8-months-of

age (e.g., Saffran, Aslin, & Newport, 1996). Indeed, a proposed computational model of early syntactic acquisition provides corroborating evidence that children might initially form syntactic "templates" based on distributional analyses of linguistic input (Cartwright & Brent, 1997). Children thus play an active role in language learning, as they come to the world equipped with powerful cognitive abilities with which to extract or induce patterns from specific instances of language use heard in their input (Akhtar, 1999; Goldberg, 1998; Tomasello, 2000a). By allowing that children are better learners than has been previously assumed (and thus that they have the mechanisms with which to learn complex structures), the need for experience-independent linguistic structures and mechanisms (i.e., linguistic parameters) is obviated.

Can a Discourse-Pragmatic Account Explain the Development of Argument Representation?

The main impetus for having explored children's learning of argument representation from a discourse-pragmatic perspective was based on the assumption that discourse-pragmatic strategies, in being available to children from very early on, would provide continuity of processes from infancy to adulthood. The performance-based account, although it inherently assumes the continuity of processes (i.e., working memory or processing limitations are present for both children and adults), could not be extended to explaining argument representation in languages where argument omission is a grammatical option. As for the competence-based account, the lack of continuity is one of its biggest drawbacks. By assuming the continuity of processes, there is no need to postulate a more complex learning mechanism which would allow for the "unlearning" of

early argument omission (as is typically predicted under a competence-based account). Secondly, if discourse-pragmatic strategies underlying referential choices indeed stem from underlying cognitive processes and constraints, it obviates the need to postulate internal, specifically language-based, constraints (e.g., linguistic parameters).

Based on the assumption that discourse-pragmatic principles are derived from underlying cognitive processes and constraints, it was predicted that children's referential choices would show sensitivity to discourse-pragmatic constraints from very early on in language development. However, the findings obtained here indicate that children do not come to the world equipped with an implicit understanding of discourse-pragmatic principles. For instance, the English-speaking and Japanese-speaking children did not know how to lexicalize new information from the onset of speech production. It turns out that appropriate referential principles, such as the lexicalization of new information and non-lexicalization of given information, need to be learned.

As a result of these findings, it seems that the discourse-pragmatic account, on its own, cannot explain all of children's referential choices. Children's patterns of argument omission and production, in particular more idiosyncratic uses, are better explained by appealing to the influence of parental input. Indeed, the input analyses have shown that children's learning of language-specific referential strategies were due to patterns observed in parental input, indicating the important contribution of environmental input to language learning. Parental input patterns also influenced the children's use of presumed universal discourse-pragmatic principles, resulting in crosslinguistic variability between English and Japanese.

Albeit the unexpected findings, the data lend themselves well to a discourse-pragmatic account of argument representation. By having adduced evidence of discourse-pragmatic strategies in child English, the present thesis has shown that the discourse-pragmatic approach can be applied toward explaining argument representation crosslinguistically, whether the language is an overt or a null argument one, as well as whether the morphological form of the argument appears as null, pronominal, or lexical. Such evidence was previously lacking in the literature. Furthermore, the developmental analyses indicate that a discourse-pragmatic account can be adequately applied towards explaining argument choice across developmental periods. For instance, even though the English-speaking children no longer used argument omission, the strategies for the referencing of given information continue to be the same, except that the resulting arguments took the form of pronominal rather than null forms. Thus, whether learning a null or an overt argument language, the same discourse-pragmatic strategies are employed in the selection of an argument's form.

Once leeway is made for the role of input, the discourse-pragmatic account fares much better than either of the competence- or performance-based accounts. Moreover, the discourse-pragmatic account inherently assumes a cognitive, or performance, component, such as working memory or processing limitations. As suggested by those working in this field (e.g., Chafe, 1987; 1994; 1996; Du Bois, 1987), limited cognitive resources are likely reflected in the discourse conventions observed, such as the avoidance of "heavy" subjects or the tendency to avoid more than one new or lexical argument per clause. Thus, the finding that discourse-pragmatic strategies are not used by children from the beginning of speech production does not preclude the continuity of

processes, nor the relation between discourse-pragmatic strategies and underlying cognitive constraints.

In conclusion, the goal of this thesis was to investigate children's learning of argument representation. Based on the studies reported here, I have shown that the learning of argument representation comes about from the interaction of both environmental input and the general cognitive processes that the child brings to the task. Young children are capable of cognitively-sophisticated learning with which to extract both linguistic and extra-linguistic (i.e., pragmatic) patterns; patterns which are regularly provided in parental input. Once learned, argument choices stem from the interrelationship between language-universal discourse-pragmatic principles and language-specific referential strategies. Finally, by having conducted a discourse-pragmatic analysis of children's referential choices, I have shown that the learning and use of language cannot proceed without taking the social and pragmatic context into account.

Directions for Future Research

Future research is directed at investigating some of the unresolved issues brought up by the findings obtained in this thesis. One issue has to do with the role of imperatives in parental speech to children. The findings obtained in the last paper (i.e., Chapter 4) indicate that young children are exposed to a large number of imperative utterances at early stages of language development. As such, imperative structures provide children with abundant examples of the morphological representation and syntactic distribution of sentence arguments, and null arguments in particular. Could children's argument

omission in early language be the result of the frequent use of imperatives in children's input? If in fact children's first verbs are learned on an item-based manner (e.g., Tomasello, 1992a; 2000b), we would expect to see the same structures replicated in the children's speech. Future research will thus investigate the relationship between parental use of imperatives and children's use of null arguments, as well as whether extensive use of imperatives affects children's learning of language-specific argument representation.

In the second paper (i.e., Chapter 3) it was discovered that the Japanese-speaking children and mothers made inconsistent use of gestures when using null arguments to represent new information. On the other hand, the English speakers were considerably more consistent in using gestures to clarify new referents that were represented with a less specific argument form. It was suggested that the Japanese speakers omitted to use gestures precisely because the argument was omitted. Since gestures encode information that is simultaneously also expressed in speech, a language that does not verbally encode a particular referential concept would similarly not produce the corresponding gestural expression (Kita & Özyürek, 2003). However, this issue remains unresolved at the moment. Future research will more closely examine the relationship between argument representation and gesture use. It is expected that crosslinguistic differences in argument representation will reflect crosslinguistic differences in gesture use, and this will be investigated as well.

Another issue concerns the relationship between the animacy of discourse referents and their location in syntax. Findings obtained in the third paper (i.e., Chapter 4) showed that the Japanese-speaking mothers (and likely their children as well) tended to cast new arguments in intransitive subject position, as opposed to the expected transitive

object position. It has been suggested that this might occur when a new human protagonist is introduced for the first time. Due to the "animacy hierarchy," animate referents typically appear as transitive subjects and inanimate ones appear as transitive objects (Du Bois, 1987), but because the Given A Constraint restricts the location of new arguments, new animates are cast as intransitive subjects instead. Future research will investigate whether the Japanese-speaking mothers' frequent use of new intransitive subjects might be due to the animate status of the referent. In addition, future research will also be directed towards investigating the relationship between verb transitivity and the distribution of animate versus inanimate referents in children's speech. Some research exists at present which suggests that children, too, are sensitive to the animacy hierarchy (Akhtar, 1999; Allen & Schröder, 2003; Clancy, 2003; Dodson & Tomasello, 1998; Lempert, 1989). It has been suggested that children might have a general cognitive understanding that animate agents act on inanimate patients (Tomasello, 1992b; 2000a), which bootstraps them into the learning of the transitive structure. However, the role of parental input in children's learning of this pattern remains unresolved.

A related unresolved issue has to do with the finding from the third study that the English-speaking and Japanese-speaking mothers showed more frequent use of lexical arguments in transitive subject position than is generally reported in the literature (e.g., Du Bois, 1987). Closer inspection of the data revealed that the lexical arguments reflected the use of proper names or kinship terms in self- or addressee-reference (e.g., *Mommy will help you with that*, where a kinship term is used in place of a first-person pronoun). Preliminary analyses suggest that lexical arguments used as transitive subjects do in fact represent given animate referents (Guerriero, Oshima-Takane, & Ono, 2003),

and future research will explore this in more detail. Likewise, it was observed that the English-speaking mothers and children tended to cast new arguments in transitive subject role more frequently than is generally reported (Allen & Schröder, 2003; Clancy, 2003; Du Bois, 1987). Closer inspection indicated that these uses of new transitive subjects represented animate third person referents, usually the fathers, who were not present during the interaction. Future research will thus investigate the relationship between argument representation, referent animacy, and the differential uses of proper names (or kinship terms) and personal pronouns, as well as the influence of parental input on children's learning of this relationship.

References

Akhtar, N. (1999). Acquiring basic word order: Evidence for data-driven learning of syntactic structure. *Journal of Child Language*, *26*, 339-356.

Allen, S. E. M. (1997). A discourse-pragmatic explanation for the subject-object asymmetry in early null arguments: The Principle of Informativeness revisited. In A. Sorace, C. Heycock, & R. Shillcock (Eds.), *Proceedings of the GALA Conference on Language Acquisition* (pp. 10-15). Edinburgh, Scotland: Human Communication Research Centre.

Allen, S. E. M. (1999, July). *Discourse-pragmatics as an explanation for the distribution of forms in arguments*. Paper presented at the VIIIth Congress of the International Association for the Study of Child Language. San Sebastian-Donostia, Basque Country, Spain.

Allen, S. E. M. (2000). A discourse-pragmatic explanation for argument representation in child Inuktitut. *Linguistics*, *38*, 483-521.

Allen, S. E. M. & Schröder, H. (2003). Preferred argument structure in early Inuktitut spontaneous speech data. In S. A. Thompson & P. J. Hopper (Series Eds.) & J. W. Du Bois, L. E. Kumpf, & W. J. Ashby (Vol. Eds.), *Studies in discourse and grammar: Vol. 14. Preferred Argument Structure: Grammar as architecture for function* (pp. 301-338). Amsterdam: John Benjamins Publishing.

Ariel, M. (1990). *Accessing Noun-Phrase antecedents*. London, UK: Routledge.

Ariel, M. (1996). Referring expressions and the +/– coreference distinction. In J. L. Mey, H. Parret, & J. Verschueren (Series Eds.) & T. Fretheim & J. K. Gundel (Vol.

Eds.), *Pragmatics and beyond new series: Vol. 38. Reference and referent accessibility* (pp. 13-36). Amsterdam: John Benjamins Publishing.

Ashby, W. J. & Bentivoglio, P. (1993). Preferred argument structure in spoken French and Spanish. *Language Variation and Change, 5*, 61-76.

Austin, J., Blume, M., Parkinson, D., Núñez del Prado, Z., & Lust, B. (1997a). The status of *pro*-drop in the initial state: Results from new analyses of Spanish. In A. T. Pérez-Leroux & W. R. Glass (Eds.), *Contemporary perspectives on the acquisition of Spanish: Vol. 1. Developing grammars* (pp. 37-54). Somerville, MA: Cascadilla Press.

Austin, J., Blume, M., Parkinson, D., Núñez del Prado, Z., & Lust, B. (1997b). The status of *pro*-drop in the initial state: Results from new analyses of Spanish. In A. T. Pérez-Leroux & W. R. Glass (Eds.), *Contemporary perspectives on the acquisition of Spanish: Vol. 1. Developing grammars* (pp. 37-54). Somerville, MA: Cascadilla Press.

Baker, N. D. & Greenfield, P. M. (1988). The development of new and old information in young children's early language. *Language Sciences, 10*, 3-34.

Bard, E. G., Anderson, A. H., Sotillo, C., Aylett, M., Doherty-Sneddon, G., & Newlands, A. (2000). Controlling the intelligibility of referring expressions in dialogue. *Journal of Memory and Language, 42*, 1-22.

Barner, D., Guerriero, A. M. S., & Oshima-Takane, Y. (2001). The calculation of Mean Length of Utterance: Methods for assigning transcripts to Brown's linguistic stages. In H. Sirai (Ed.), *A crosslinguistic study for the Universal Developmental Index* (pp. 223-244). Chukyo, Japan: Chukyo University.

Bates, E. (1976). *Language and context: The acquisition of pragmatics*. New York: Academic Press.

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

Bloom, L., Miller, P., & Hood, L. (1975). Variation and reduction as aspects of competence in language development. In A. D. Pick (Vol. Ed.), *Minnesota Symposia on Child Psychology* (, Vol. 9, pp. 3-55). Minneapolis, MN: The University of Minnesota Press.

Bloom, P. (1990). Subjectless sentences in child language. *Linguistic Inquiry, 21*, 491-504.

Bloom, P. (1993). Grammatical continuity in language development: The case of subjectless sentences. *Linguistic Inquiry, 24*, 721-734.

Braine, M. D. S. (1963). The ontogeny of English phrase structure: The first phrase. *Language, 39*, 1-13.

Braine, M. D. S. (1974). Length constraints, reduction rules, and holophrastic processes in children's word combinations. *Journal of Verbal Learning and Verbal Behavior, 13*, 448-456.

Bromberg, H. S. & Wexler, K. (1995). Null subjects in child *Wh*-Questions. *MIT Working Papers in Linguistics, 26*, 221-247.

Brown, R. (1973). *A first language: The early stages*. Cambridge, MA: Harvard University Press.

Brown, R. & Fraser, C. (1963). The acquisition of syntax. In C. N. Cofer & B. S. Musgrave (Vol. Eds.), *McGraw-Hill series in psychology: Verbal behavior and learning: Problems and processes* (pp. 158-197). New York: McGraw-Hill.

Butcher, C. & Goldin-Meadow, S. (2000). Gesture and the transition from one- to two-word speech: When hand and mouth come together. In S. C. Levinson (Series Ed.) & D. McNeill (Vol. Ed.), *Language, culture, and cognition: Vol.2. Language and gesture* (pp. 235-257). Cambridge, UK: Cambridge University Press.

Cameron-Faulkner, T., Lieven, E. V. M., & Tomasello, M. (2003). A construction based analysis of child directed speech. *Cognitive Science, 27*, 843-873.

Campbell, A. L., Brooks, P., & Tomasello, M. (2000). Factors affecting young children's use of pronouns as referring expressions. *Journal of Speech, Language, and Hearing Research, 43*, 1337-1349.

Cartwright, T. A. & Brent, M. R. (1997). Syntactic categorization in early language acquisition: Formalizing the role of distributional analysis. *Cognition, 63*, 121-170.

Chafe, W. L. (1987). Cognitive constraints on information flow. In R. S. Tomlin (Vol. Ed.), *Typological studies in language: Vol. 11. Coherence and grounding in discourse* (pp. 21-51). Amsterdam: John Benjamins Publishing.

Chafe, W. L. (1994). *Discourse, consciousness, and time: The flow and displacement of conscious experience in speaking and writing.* Chicago: The University of Chicago Press.

Chafe, W. L. (1996). Inferring identifiability and accessibility. In J. L. Mey, H. Parret, & J. Verschueren (Series Eds.) & T. Fretheim & J. K. Gundel (Vol. Eds.),

Pragmatics and beyond new series: Vol. 38. Reference and referent accessibility (pp. 37-46). Amsterdam: John Benjamins Publishing.

Choi, S. & Gopnik, A. (1995). Early acquisition of verbs in Korean: A crosslinguistic study. *Journal of Child Language, 22*, 497-529.

Clahsen, H. (1986). Verb inflections in German child language: Acquisition of agreement markings and the functions they encode. *Linguistics, 24*, 79-121.

Clahsen, H. (1991). Constraints on parameter setting: A grammatical analysis of some acquisition stages in German child language. *Language Acquisition, 1*, 361-391.

Clancy, P. M. (1986a). The acquisition of communicative style in Japanese. In B. B. Schieffelin & E. Ochs (Vol. Eds.), *Studies in the social and cultural foundations of language: Vol. 3. Language socialization across cultures* (pp. 213-250). Cambridge, UK: Cambridge University Press.

Clancy, P. M. (1986b). The acquisition of Japanese. In D. I. Slobin (Ed.), *The crosslinguistic study of language acquisition* (Vol. 1, pp. 373-524). Mahwah, NJ: Lawrence Erlbaum Associates.

Clancy, P. M. (1993). Preferred argument structure in Korean acquisition. In E. V. Clark (Ed.), *Proceedings of the twenty-fifth annual Child Language Research Forum* (pp. 307-314). Stanford, CA: CSLI Publications.

Clancy, P. M. (1997). Discourse motivations for referential choice in Korean acquisition. *Japanese/Korean Linguistics, 6*, 639-659.

Clancy, P. M. (2003). The lexicon in interaction: Developmental origins of preferred argument structure in Korean. In S. A. Thompson & P. J. Hopper (Series Eds.) & J. W. Du Bois, L. E. Kumpf, & W. J. Ashby (Vol. Eds.), *Studies in discourse and*

grammar: Vol. 14. Preferred Argument Structure: Grammar as architecture for function (pp. 81-108). Amsterdam: John Benjamins Publishing.

Clark, H. H. & Haviland, S. E. (1977). Comprehension and the given-new contract. In R. O. Freedle (Vol. Ed.), *Discourse production and comprehension: Vol. I. Discourse processes: Advances in research and theory* (pp. 1-40). Norwood, NJ: Ablex Publishing Corporation.

de Haan, G. J. & Tuijnman, K. (1988). Missing subjects and objects in child grammar. In P. Jordens & J. Lalleman (Eds.), *Language development* (pp. 101-121). Dordrecht, The Netherlands: Foris Publications.

Demuth, K., Machobane, 'M., & Moloi, F. (2000). Learning word-order constraints under conditions of object ellipsis. *Linguistics*, *38*, 545-568.

Dodson, K. & Tomasello, M. (1998). Acquiring the transitive construction in English: The role of animacy and pronouns. *Journal of Child Language*, *25*, 605-622.

Du Bois, J. W. (1987). The discourse basis of ergativity. *Language*, *63*, 805-855.

Du Bois, J. W. (2003). Argument structure. In S. A. Thompson & P. J. Hopper (Series Eds.) & J. W. Du Bois, L. E. Kumpf, & W. J. Ashby (Vol. Eds.), *Studies in discourse and grammar: Vol. 14. Preferred Argument Structure: Grammar as architecture for function* (pp. 11-60). Amsterdam: John Benjamins Publishing.

Ervin, S. M. (1964). Imitation and structural change in children's language. In E. H. Lenneberg (Ed.), *New directions in the study of language* (pp. 163-189). Cambridge, MA: The MIT Press.

Feldman, H., Goldin-Meadow, S., & Gleitman, L. R. (1978). Beyond Herodotus: The creation of language by linguistically deprived deaf children. In A. Lock & J. S.

Bruner (Eds.), *Action, gesture, and symbol: The emergence of language* (pp. 351-414). New York: Academic Press.

Ferguson, G. A. & Takane, Y. (1989). *Statistical analysis in psychology and education* (6th ed.). New York: McGraw-Hill Book Company.

Fisher, C. L. & Tokura, H. (1995). The given-new contract in speech to infants. *Journal of Memory and Language, 34*, 287-310.

Gerken, L. (1991). The metrical basis for children's subjectless sentences. *Journal of Memory and Language, 30*, 431-451.

Givón, T. (1983). Topic continuity in discourse: An introduction. In J. H. Greenberg & T. Givón (Series Eds.) & T. Givón (Vol. Ed.), *Typological studies in language: Vol. 3. Topic continuity in discourse: A quantitative cross-language study* (pp. 1-41). Amsterdam: John Benjamins Publishing.

Goldberg, A. E. (1998). Patterns of experience in patterns of language. In M. Tomasello (Ed.), *The new psychology of language: Cognitive and functional approaches to language structure* (Vol. 1, pp. 203-219). Mahwah, NJ: Lawrence Erlbaum Associates.

Greenfield, P. M. (1979). Informativeness, presupposition, and semantic choice in single-word utterances. In E. Ochs & B. B. Schieffelin (Eds.), *Developmental pragmatics* (pp. 159-166). New York: Academic Press.

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

Greenfield, P. M. & Zukow, P. G. (1978). Why do children say what they say when they say it?: An experimental approach to the psychogenesis of presupposition. In K.

E. Nelson (Ed.), *Children's language* (Vol. 1, pp. 287-336). New York: Gardner Press.

Guerriero, A. M. S., Cooper, A., Oshima-Takane, Y., & Kuriyama, Y. (2001). A discourse-pragmatic explanation for argument realization and omission in English and Japanese children's speech. *Proceedings of the Annual Boston University Conference on Language Development, 25*, 319-330.

Guerriero, A. M. S. & Oshima-Takane, Y. (2004). *Children's referential choices: The role of non-linguistic pragmatic information.* Manuscript in preparation, McGill University, Montréal, Canada.

Guerriero, A. M. S., Oshima-Takane, Y., Cooper, A., Kayama, Y., & Ono, A. (2001). *Null Argument Project: Manual for the coding of morphological form, syntax, and pragmatics*. McGill University. Montréal, Canada.

Guerriero, A. M. S., Oshima-Takane, Y., Cooper, A., & Kuriyama, Y. (2001, April). *Discourse-pragmatic strategies for argument choice in English and Japanese: Does children's speech reflect input?* Poster presented at the biennial meeting of the Society for Research in Child Development. Minneapolis, MN.

Guerriero, A. M. S., Oshima-Takane, Y., & Kuriyama, Y. (2004). *The development of argument representation in English and Japanese child language: A discourse-pragmatic perspective*. Manuscript submitted for publication, McGill University, Montréal, Canada.

Guerriero, A. M. S., Oshima-Takane, Y., & Ono, A. (2002a, August). *The development of argument representation in English and Japanese: Does children's speech reflect*

input? Poster presented at the 17th biennial meeting of the International Society for the Study of Behavioural Development. Ottawa, Ontario, Canada.

Guerriero, A. M. S., Oshima-Takane, Y., & Ono, A. (2002b, July). *The development of argument representation: A crosslinguistic study of English and Japanese*. Poster presented at the joint conference of the IX International Congress for the Study of Child Language and the 23rd annual Symposium on Research in Child Language Disorders. Madison, WI.

Guerriero, A. M. S., Oshima-Takane, Y., & Ono, A. (2003, April). *Animacy and discourse-pragmatic strategies: An investigation of Japanese- and English-speaking children and their mothers*. Poster presented at the biennial meeting of the Society for Research in Child Development. Tampa, FL.

Guidetti, M. (2002). The emergence of pragmatics: Forms and functions of conventional gestures in young French children. *First Language*, *22*, 285.

Guilfoyle, E. (1984). The acquisition of tense and the emergence of lexical subjects in child grammars of English. *McGill Working Papers in Linguistics*, *2*, 20-30.

Gundel, J. K. (1996). Relevance theory meets the givenness hierarchy: An account of inferrables. In J. L. Mey, H. Parret, & J. Verschueren (Series Eds.) & T. Fretheim & J. K. Gundel (Vol. Eds.), *Pragmatics and beyond new series: Vol. 38. Reference and referent accessibility* (pp. 141-154). Amsterdam: John Benjamins Publishing.

Gundel, J. K., Hedberg, N., & Zacharski, R. (1993). Cognitive status and the form of referring expressions in discourse. *Language*, *69*, 274-307.

Haegeman, L. (1991). *Introduction to government and binding theory*. Oxford, UK: Blackwell Publishers.

Haegeman, L. (1996). Root infinitives, clitics, and truncated structures. In H. Clahsen & W. Rutherford (Series Eds.) & H. Clahsen (Vol. Ed.), *Language acquisition and language disorders: Vol. 14. Generative perspectives on language acquisition: Empirical findings, theoretical considerations, and crosslinguistic comparisons* (pp. 271-307). Amsterdam: John Benjamins Publishing.

Hamann, C. (1995). Null arguments in German child language. *Proceedings of the Annual Boston University Conference on Language Development, 19*, 240-254.

Hamann, C. (1996). Null arguments in German child language. *Language Acquisition, 5*, 155-208.

Hamann, C. & Plunkett, K. (1998). Subjectless sentences in child Danish. *Cognition, 69*, 35-72.

Hirakawa, M. (1993). Null subjects versus null objects in an early grammar of Japanese. *McGill Working Papers in Linguistics, 9*, 30-45.

Hoekstra, T. & Hyams, N. M. (1996). Missing heads in child language. In C. Koster & F. Wijnen (Eds.), *Proceedings of the Groningen Assembly on Language Acquisition* (pp. 251-260). Groningen, The Netherlands: Centre Language and Cognition.

Hoff-Ginsberg, E. (1985). Some contributions of mothers' speech to their children's syntactic growth. *Journal of Child Language, 12*, 367-385.

Hyams, N. M. (1986). *Language acquisition and the theory of parameters*. Dordrecht, The Netherlands: D. Reidel Publishing.

Hyams, N. M. (1992). A reanalysis of null subjects in child language. In J. Weissenborn, H. Goodluck, & T. Roeper (Eds.), *Theoretical issues in language acquisition: Continuity and change in development* (pp. 249-267). Hillsdale, NJ: Lawrence Erlbaum Associates.

Hyams, N. M. (1996). The underspecification of functional categories in early grammar. In H. Clahsen & W. Rutherford (Series Eds.) & H. Clahsen (Vol. Ed.), *Language acquisition and language disorders: Vol. 14. Generative perspectives on language acquisition: Empirical findings, theoretical considerations, and crosslinguistic comparisons* (pp. 91-127). Amsterdam: John Benjamins Publishing.

Hyams, N. M. (2001). Now you hear it, now you don't: The nature of optionality in child grammars. *Proceedings of the Annual Boston University Conference on Language Development, 25*, 34-58.

Hyams, N. M. & Wexler, K. (1993). On the grammatical basis of null subjects in child language. *Linguistic Inquiry, 24*, 421-459.

Ingham, R. (1992). The optional subject phenomenon in young children's English: A case study. *Journal of Child Language, 19*, 133-151.

Jaeggli, O. A. & Hyams, N. M. (1987). Morphological uniformity and the setting of the null subject parameter. *Proceedings of the North East Linguistic Society, 18*, 238-253.

Kärkkäinen, E. (1996). Preferred argument structure and subject role in American English conversational discourse. *Journal of Pragmatics, 25*, 675-701.

Kim, Y. (1997). The acquisition of Korean. In D. I. Slobin (Ed.), *The crosslinguistic study of language acquisition: Volume 4* (Vol. 4, pp. 335-443). Mahwah, NJ: Lawrence Erlbaum Associates.

Kim, Y.-J. (2000). Subject/object drop in the acquisition of Korean: A crosslinguistic comparison. *Journal of East Asian Linguistics*, *9*, 325-351.

Kita, S. & Özyürek, A. (2003). What does cross-linguistic variation in semantic coordination of speech and gesture reveal?: Evidence for an interface representation of spatial thinking and speaking. *Journal of Memory and Language*, *48*, 16-32.

Kumpf, L. E. (1992). Preferred Argument Structure in second language discourse: A preliminary study. *Studies in Language*, *16*, 369-403.

Kuriyama, Y. (1999). The analysis of Japanese predicate verbs in utterances by children aged 36 months. *IX European Conference in Developmental Psychology Abstracts* 78.

Lambrecht, K. (1994). *Information structure and sentence form: Topic, focus, and the mental representation of discourse referents*. Cambridge, UK: Cambridge University Press.

Lempert, H. (1989). Animacy constraints on preschool children's acquisition of syntax. *Child Development*, *60*, 237-245.

Lieven, E. V. M., Pine, J. M., & Baldwin, G. (1997). Lexically-based learning and early grammatical development. *Journal of Child Language*, *24*, 187-219.

Lillo-Martin, D. (1986a). Effects of the acquisition of morphology on syntactic parameter setting. *Proceedings of the North East Linguistic Society*, *16*, 305-321.

Lillo-Martin, D. (1986b). Two kinds of null arguments in American Sign Language. *Natural Language and Linguistic Theory, 4*, 415-444.

Lillo-Martin, D. (1994). Setting the null argument parameters: Evidence from American Sign Language and other languages. In B. Lust, G. Hermon, & J. Kornfilt (Eds.), *Syntactic theory and first language acquisition: Cross-linguistic perspectives: Vol. 2. Binding, dependencies, and learnability* (pp. 301-318). Hillsdale, NJ: Lawrence Erlbaum Associates.

MacWhinney, B. (2000a). *The CHILDES project: Tools for analyzing talk: Vol. 1. The format and programs* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

MacWhinney, B. (2000b). *The CHILDES project: Tools for analyzing talk: Vol. 2. The database* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

Matsumoto, K. (2000). Intonation units, clauses, and preferred argument structure in conversational Japanese. *Language Sciences, 22*, 63-86.

Mayberry, R. I. & Jaques, J. (2000). Gesture production during stuttered speech: Insights into the nature of gesture-speech integration. In S. C. Levinson (Series Ed.) & D. McNeill (Vol. Ed.), *Language, culture, and cognition: Vol.2. Language and gesture* (pp. 199-214). Cambridge, UK: Cambridge University Press.

Mazuka, R., Lust, B., Wakayama, T., & Snyder, W. (1986). Distinguishing effects of parameters in early syntax acquisition: A cross-linguistic study of Japanese and English. *Papers and Reports on Child Language Development, 25*, 73-82.

McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. Chicago: The University of Chicago Press.

Menyuk, P. (1969). *Sentences children use*. Cambridge, MA: The MIT Press.

Miyata, S. (1995). The Aki corpus: Longitudinal speech data of a Japanese boy aged 1.6-2.12. *Bulletin of Aichi Shukutoku Junior College*, *34*, 183-191.

Miyata, S. (2000). The Tai corpus: Longitudinal speech data of a Japanese boy aged 1;5.20-3;1.1. *Bulletin of Aichi Shukutoku Junior College*, *39*, 77-85.

Miyata, S., Oshima-Takane, Y., & Nisisawa, H. Y. (2004). Noun bias in early Japanese vocabulary and characteristics of maternal speech. *Studies in Language Sciences*, *3*, 87-101.

Naigles, L. R. (2002). Form is easy, meaning is hard: Resolving a paradox in early child language. *Cognition*, *86*, 157-199.

Naigles, L. R. & Hoff-Ginsberg, E. (1995). Input to verb learning: Evidence for the plausibility of syntactic bootstrapping. *Developmental Psychology*, *31*, 827-837.

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

Nakayama, M. (1994). Null arguments in Japanese children's speech. *MIT Working Papers in Linguistics*, *24*, 247-261.

Nooteboom, S. G. & Kruyt, J. G. (1987). Accents, focus distribution, and the perceived distribution of given and new information: An experiment. *Journal of the Acoustical Society of America*, *85*, 1512-1524.

O'Grady, W., Peters, A. M., & Masterson, D. (1989). The transition from optional to required subjects. *Journal of Child Language*, *16*, 513-529.

O'Neill, D. K. & Topolovec, J. C. (2001). Two-year-old children's sensitivity to the referential (in)efficiency of their own pointing gestures. *Journal of Child Language, 28*, 1-28.

Oshima-Takane, Y. Oshima-Takane, Y., MacWhinney, B., Sirai, H., Miyata, S., & Naka, N. (Eds.). (1998). *Nihongo no tame no CHILDES manyuaru [CHILDES manual for Japanese]* (2nd ed.). Nagoya, Japan: JCHAT Project.

Oshima-Takane, Y. (in press). Acquisition of nouns and verbs in Japanese. *Handbook of Japanese Psycholinguistics.*

Oshima-Takane, Y., Goodz, E., & Derevensky, J. L. (1996). Birth order effects on early language development: Do secondborn children learn from overheard speech? *Child Development, 67*, 621-634.

Oshima-Takane, Y., MacWhinney, B., Sirai, H., Miyata, S., & Naka, N. (1998). *(Eds.). Nihongo no tame no CHILDES manyuaru [CHILDES manual for Japanese]* (2nd ed.). Nagoya, Japan: JCHAT Project.

Oshima-Takane, Y. & Oram, J. (1991, July). *Acquisition of personal pronouns: What do comprehension data tell us?* Poster presented at the biennial meeting of the International Society for the Study of Behavioural Development. Minneapolis, MN.

Pechmann, T. & Deutsch, W. (1982). The development of verbal and nonverbal devices for reference. *Journal of Experimental Child Psychology, 34*, 330-341.

Pierce, A. (1994). On the differing status of subject pronouns in French and English child language. In B. Lust, G. Hermon, & J. Kornfilt (Eds.), *Syntactic theory and first language acquisition: Cross-linguistic perspectives: Vol. 2. Binding,*

dependencies, and learnability (pp. 319-333). Hillsdale, NJ: Lawrence Erlbaum Associates.

Pinker, S. (1984). *Language learnability and language development.* Cambridge, MA: Harvard University Press.

Prince, E. F. (1981). Toward a taxonomy of given-new information. In P. Cole (Ed.), *Radical Pragmatics* (pp. 223-255). New York: Academic Press.

Radford, A. (1990). *Syntactic theory and the acquisition of English syntax: The nature of early child grammars of English.* Oxford, UK: Basil Blackwell.

Radford, A. (1996). Towards a structure-building model of acquisition. In H. Clahsen & W. Rutherford (Series Eds.) & H. Clahsen (Vol. Ed.), *Language acquisition and language disorders: Vol. 14. Generative perspectives on language acquisition: Empirical findings, theoretical considerations, and crosslinguistic findings* (pp. 43-89). Amsterdam: John Benjamins Publishing.

Rizzi, L. (1994). Early null subjects and root null subjects. In B. Lust, G. Hermon, & J. Kornfilt (Eds.), *Syntactic theory and first language acquisition: Cross-linguistic perspectives: Vol. 2. Binding, dependencies, and learnability* (pp. 249-272). Hillsdale, NJ: Lawrence Erlbaum Associates.

Rowland, C. F., Pine, J. M., Lieven, E. V. M., & Theakston, A. L. (2003). Determinants of acquisition order in wh-questions: Re-evaluating the role of caregiver speech. *Journal of Child Language, 30,* 609-635.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science, 274,* 1926-1928.

Sano, T. & Hyams, N. M. (1994). Agreement, finiteness, and the development of null arguments. *Proceedings of the North East Linguistic Society, 24*, 543-558.

Schütze, C. T. & Wexler, K. (1996). Subject case licensing and English root infinitives. *Proceedings of the Annual Boston University Conference on Language Development, 20*, 670-681.

Sigurjónsdóttir, S. (1999). Root infinitives and null subjects in early Icelandic. *Proceedings of the Annual Boston University Conference on Language Development, 23*, 630-641.

Skarabela, B. & Allen, S. E. M. (2002). The role of joint attention in argument realization in child Inuktitut. *Proceedings of the Annual Boston University Conference on Language Development, 26*, 620-630.

Sonnenschein, S. (1986). Development of referential communication skills: How familiarity with a listener affects a speaker's production of redundant messages. *Developmental Psychology, 22*, 549-552.

Strömqvist, S. & Ragnarsdóttir, H. (2000). On the acquisition of verb argument structure. *Linguistics, 38*, 523-543.

Tardif, T., Shatz, M., & Naigles, L. R. (1997). Caregiver speech and children's use of nouns versus verbs: A comparison of English, Italian, and Mandarin. *Journal of Child Language, 24*, 535-565.

Terken, J. & Hirschberg, J. (1994). Deaccentuation of words representing 'given' information: Effects of persistence of grammatical function and surface position. *Language and Speech, 37*, 125-145.

Theakston, A. L., Lieven, E. V. M., Pine, J. M., & Rowland, C. F. (2001). The role of performance limitations in the acquisition of verb-argument structure: An alternative account. *Journal of Child Language, 28*, 127-152.

Theakston, A. L., Lieven, E. V. M., Pine, J. M., & Rowland, C. F. (2004). Semantic generality, input frequency and the acquisition of syntax. *Journal of Child Language, 31*, 61-99.

Thoermer, C. & Sodian, B. (2001). Preverbal infants' understanding of referential gestures. *First Language, 21*, 245-264.

Tomasello, M. (1992a). *First verbs: A case study of early grammatical development.* Cambridge, UK: Cambridge University Press.

Tomasello, M. (1992b). The social bases of language acquisition. *Social Development, 1*, 67-87.

Tomasello, M. (2000a). Do young children have adult syntactic competence? *Cognition, 74*, 209-253.

Tomasello, M. (2000b). First steps toward a usage-based theory of language acquisition. *Cognitive Linguistics, 11*, 61-82.

Tomasello, M., Anselmi, D., & Farrar, M. J. (1985). Young children's coordination of gestural and linguistic reference. *First Language, 5*, 199-210.

Tomasello, M. & Haberl, K. (2003). Understanding attention: 12- and 18-month-olds know what is new for other persons. *Developmental Psychology, 39*, 906-912.

Uziel-Karl, S. & Berman, R. A. (2000). Where's ellipsis? Whether and why there are missing arguments in Hebrew child language. *Linguistics, 38*, 457-482.

Valian, V. (1990). Null subjects: A problem for parameter-setting models of language acquisition. *Cognition, 35*, 105-122.

Valian, V. (1991). Syntactic subjects in the early speech of American and Italian children. *Cognition, 40*, 21-81.

Valian, V. (1994). Children's postulation of null subjects: Parameter setting and language acquisition. In B. Lust, G. Hermon, & J. Kornfilt (Eds.), *Syntactic theory and first language acquisition: Cross-linguistic perspectives: Vol. 2. Binding, dependencies, and learnability* (pp. 273-286). Hillsdale, NJ: Lawrence Erlbaum Associates.

Valian, V. & Eisenberg, Z. (1996). The development of syntactic subjects in Portuguese-speaking children. *Journal of Child Language, 23*, 103-128.

Vygotsky, L. S. (1962). *Thought and language*. Cambridge, MA: The MIT Press.

Wang, Q., Lillo-Martin, D., Best, C. T., & Levitt, A. (1992). Null subject versus null object: Some evidence from the acquisition of Chinese and English. *Language Acquisition, 2*, 221-254.

Weissenborn, J. (1992). Null subjects in early grammars: Implications for parameter-setting theories. In J. Weissenborn, H. Goodluck, & T. Roeper (Eds.), *Theoretical issues in language acquisition: Continuity and change in development* (pp. 269-299). Hillsdale, NJ: Lawrence Erlbaum Associates.

Wexler, K. (1994). Optional infinitives, head movement, and the economy of derivations in child grammar. In D. Lightfoot & N. Hornstein (Eds.), *Verb movement* (pp. 305-350). Cambridge, MA: Cambridge University Press.

Wilcox, M. J. & Howse, P. (1982). Children's use of gestural and verbal behavior in communicative misunderstandings. *Applied Psycholinguistics*, *3*, 15-27.

Who will the participants be?

see attached

How will participants be recruited? (Attach copies of all written or spoken material that will be used in recruiting subjects, such as newspaper ads, posted notices, and verbal announcements.)

see attached

How will organizational/community/governmental permission be obtained (if applicable)?

see attached

How will data be collected, i.e., what will the participants be asked to do?

see attached

Does the study pose any risks to participants? If so, please state why these are necessary and explain how you plan to deal with them.

see attached

Does the study involve deception? If so, please state why this is necessary and explain how you plan to deal with potential negative effects (e.g., by post-experimental debriefing).

see attached

How will you document informed consent to participate in the study? (Attach written informed consent form. If written consent is not possible, how will you document verbal consent? If it is not possible to obtain informed consent, explain why this the case.)

see attached

How will participants be informed of their right to withdraw at any time?

see attached

). How will subject/data anonymity and confidentiality be maintained?

see attached

I. Please comment on any other potential ethical concerns which may arise in the course of the research. If the proposed research involves testing subjects in situations where particular problems might arise, please explain how researchers will be trained to handle matters in a sensitive and professional way.

see attached

McGill University Research Ethics Board II

Ethical Considerations of Proposed Research Involving Humans

1. *Briefly describe the research topic.*
   This project is part of an international collaborative research project investigating how children acquire fundamental lexical categories, such as nouns and verbs. By studying lexical development in English- and Japanese-speaking children, the following issues are specifically examined in this project: (1) whether caregivers' input contains sufficient semantic as well as morphosyntactic information to enable children to distinguish fundamental lexical categories across languages and (2) to what extent children are capable of extracting and using such information from input in acquiring these categories.

2. *Who will the participants be?*
   The participants will be caregivers and their young infants who are acquiring English as their first language. Children must be between 8 and 24 months of age at the start of the study.

3. *How will participants be recruited? (Attach copies of all written or spoken material that will be used in recruiting subjects, such as newspaper ads, posted notices, and verbal announcements.)*
   A letter describing the research project and search for participants (see attachment) will be distributed to households within predominantly English-speaking regions of Montréal such as Hampstead, Cote-St.-Luc, Westmount.

4. *How will organizational/community/governmental permission be obtained (if applicable)?*
   Not applicable.

5. *How will data be collected, i.e., what will the participants be asked to do?*
   English spontaneous production data from mother-child pairs will be collected comparable to the Japanese data collected by researchers in Japan. A few toys, books, and household artifacts (a hairbrush, a spoon, a toy hammer) will be provided to the dyads and mothers' and children's spontaneous verbalizations will be observed as they interact with them in a naturalistic situation. Participants will be videotaped in their own homes for one hour on a bi-weekly basis for approximately one year, or until the child attains a Mean Length of Utterance between 3.5 and 4.0. In addition, parents will complete the MacArthur Communicative Development Inventory on a monthly basis. To obtain standardized language measures, the following standardized tests will be administered at certain sessions: the Bayley Scale of Infant Development, the Stanford-Binet, ant the Peabody Picture Vocabulary Test.

6. *Does the study pose any risks to participants? If so, please state why these are necessary and explain how you plan to deal with them.*
The study does not pose any risks to participants and there are no foreseeable physical and mental stresses to either parent or child. Parents' and children's spontaneous production data will be collected as they interact with each other in a naturalistic situation. If a child becomes distressed or uneasy during the videotaping, parents may ask to pause the session for a few minutes or to resume the entire session at another time.

7. *Does the study involve deception? If so, please state why this is necessary and explain how you plan to deal with potential negative effects (e.g., by post-experimental debriefing).*
Participants are not deceived in any way. However, to obtain as naturalistic a language sample as possible, parents will not be told that their language production will be analysed in addition to their child's. This will be disclosed to them at the last session. If, after disclosure, parents object to the analysis of their language, they may withdraw their consent to participate in the study and their data will not be used. Upon completion of the study, a written report of the findings will be made available to those parents who are interested. Parents may also request copies of the videotapes.

8. *How will you document informed consent to participate in the study? (Attach written informed consent form. If written consent is not possible, how will you document verbal consent? If it is not possible to obtain informed consent, explain why this is the case.)*
Parents will be asked to sign a consent form (see attachment).

9. *How will participants be informed of their right to withdraw at any time?*
Participants will be informed at the start of the study that they are free to discontinue at any time.

10. *How will subject/data anonymity and confidentiality be maintained?*
All data collected, including videotapes and test scores, will be kept securely in our labs. Speech and context will be transcribed from the videotapes and made accessible only to research associates from Japan and lab members, such as graduate and undergraduate students and research assistants working under the direct supervision of Yuriko Oshima-Takane. Data will be used exclusively for research purposes. Participants will be identified by pseudonyms in all published reports of the study. Once all data has been transcribed, parents will be asked for permission to donate their data to the CHILDES international database.

11. *Please comment on any other potential ethical concerns which may arise in the course of the research. If the proposed research involves testing subjects in situations where particular problems might arise, please explain how researchers will be trained to handle matters in a sensitive and professional way.*
There are no other potential ethical concerns.

McGill

A M Sonia Guerriero
.232.

Department of Psychology

Stewart Biological Sciences Building
1205 Dr. Penfield Avenue
Montreal, QC, Canada H3A 1B1

Département de psychologie

Pavillon Stewart des Sciences Biologiques
1205, avenue Dr. Penfield
Montréal, QC, Canada H3A 1B1

Tel.: (514) 398-6100
Fax: (514) 398-4896

CONSENT FORM

I, _____, (please print your full name) agree to participate in a study entitled ***A crosslinguistic study of the acquisition of lexical categories in English and Japanese***, and I give consent for my child, _____, (please print your child's full name) to also take part in this study. Our participation in this study is entirely voluntary. The study has received ethics approval from McGill University and will be administered by Sonia Guerriero, PhD student, under the supervision of Dr. Yuriko Oshima-Takane of the Department of Psychology at McGill University.

The purpose of the study is to investigate how children learn to use words, such as nouns and verbs, by observing their language as it occurs spontaneously. Participation in this study requires that the researchers videotape my child, myself, and an older-sibling as we interact with each other naturally. The videotaping will take place in my own home on a bi-weekly basis and will continue until my child has either reached three years of age or has acquired language (almost) fully. Each session will last approximately one and one half hours. I will be asked to complete the MacArthur Communicative Development Inventory and a checklist of vocabulary development on a monthly basis. At certain sessions, the following standardized tests may be given to my child: the Bayley Scale of Infant Development, the Stanford-Binet, and the Peabody Picture Vocabulary Test. The researchers who will be present at these sessions are Sonia Guerriero and one or two trained research assistants. I give permission for a trained research assistant to substitute for Sonia Guerriero as the study progresses.

The procedure described above does not involve any foreseeable risks to me or my child(ren). If my child becomes distressed or uneasy during a session or if unexpected circumstances occur, I may ask to have the session paused for a few minutes or to resume the entire session at a later time. I have the right to withdraw myself and my child(ren) at any time.

All data collected from me and my child(ren), including assessment scores and videotapes, become the property of Dr. Oshima-Takane's Developmental Psycholinguistics Laboratory. My name, the name of my child(ren), and the data collected from us will be kept confidential. Only individuals directly connected with the study will have access to the data. All data will be used for research purposes only and will be reported as part of a group in published reports of the study. If particular linguistic episodes of myself or my child(ren) are to be discussed in isolation, pseudonyms or initials will be used.

By signing below, I have read and agree to the above statements.

_____ _____

parent's signature date

_____ _____

Sonia Guerriero, Ph.D. candidate date
Yuriko Oshima-Takane, Ph.D., Project Director
Department of Psychology, McGill University

McGill

Department of Psychology

Stewart Biological Sciences Building
1205 Dr. Penfield Avenue
Montreal, QC, Canada H3A 1B1

Département de psychologie

Pavillon Stewart des Sciences Biologiques
1205, avenue Dr. Penfield
Montréal, QC, Canada H3A 1B1

Tel.: (514) 398-6100
Fax: (514) 398-4896

June 2000

Dear parent(s):

We are a team of researchers from the Department of Psychology at McGill University investigating children's language development and are recruiting participants for a new study. If you have pre-school aged children, the following study may be of interest.

This new study is a crosslinguistic investigation into how different types of words are acquired in English- and Japanese-speaking children. One issue we are investigating is how children acquire nouns and verbs and how they learn to use them in a sentence. At the moment, we are recruiting English-speaking participants.

We are looking for children from English-speaking families who are between 8 and 24 months of age. In order to conduct a study of this nature, we will videotape each child with his/her mother on a bi-weekly basis as they interact naturally in their own home. Each bi-weekly session will last approximately one hour and will continue for the duration of approximately one year. Parents will be asked to complete a vocabulary checklist on a monthly basis.

The study has received ethics approval from McGill University and parents are assured that every effort is made to protect the interest of the children. Participants are free to discontinue at any time. At the end of the project, a written report of the results will be sent to all parents who wish to receive one. To thank them for their participation, parents may request copies of the videotapes.

We invite interested parents to contact Sonia Guerriero for more specific details regarding the study. We thank you for your time and look forward to hearing from you.

Sincerely,

Yuriko Oshima-Takane, Ph.D.
Project Director
(514) 398-4672
Yuriko@hebb.psych.mcgill.ca
Department of Psychology, McGill University

Sonia Guerriero, Ph.D. candidate
(514) 398-7098
(514) 398-4455 x1291
SoniaG@ego.psych.mcgill.ca
Department of Psychology, McGill University