Information management in Mandarin child speech, maternal speech, and adult speech[☆]

Chiung-chih Huang [a,b,*]

[a] *Graduate Institute of Linguistics, National Chengchi University, No. 64, Sec. 2, Zhi-nan Rd., Wenshan District, Taipei City 11605, Taiwan*
[b] *Research Center for Mind, Brain, and Learning, National Chengchi University, No. 64, Sec. 2, Zhi-nan Rd., Wenshan District, Taipei City 11605, Taiwan*

Abstract

This study investigated information management in Mandarin child speech, maternal speech, and adult speech. The three types of speech were analyzed to test the hypothesis of preferred argument structure (Du Bois, 1987, 2003), and they were also analyzed in terms of two types of discourse information measures. The data consisted of eight hours of mother–child (aged 2;2 to 3;1) conversation and one hour of adult–adult conversation. The analysis of the data showed that the three types of speech demonstrated similar patterns in grammatical alignment and that all of them followed the constraints of quantity. The discourse information measures, on the other hand, showed that the three types of speech differed significantly in the measure of information pressure and in the measure of lexical referential density. The results revealed how information management may be associated with children's developing linguistic ability, mothers' speech adjustments, and adults' mature competence.
© 2016 Elsevier B.V. All rights reserved.

Keywords: Information management; Argument structure; Mandarin Chinese; Mother–child conversation; Language acquisition

1. Introduction

1.1. Preferred argument structure

Information management in discourse has been extensively studied in the function–oriented research of language. One of the important theories of information management is Du Bois' preferred argument structure (henceforth PAS) (1987, 2003). PAS concerns the flow of information in discourse and its interaction with the primary noun arguments associated with verb phrases: The subject of a transitive verb (A), the object of a transitive verb (O), and the subject of an intransitive verb (S).

The central notions of PAS can be expressed in the form of four constraints, which are hypothesized to be universal across languages. The four constraints include two quantity constraints and two role constraints. The two quantity

---

Table 1
Dimensions and constraints of preferred argument structure.

	Grammar	Pragmatics
Quantity	One Lexical Argument Constraint	One New Argument Constraint
Role	Non-lexical A Constraint	Given A Constraint

Source: Adapted from Du Bois (1987, p. 829).

constraints are the One New Argument Constraint and the One Lexical Argument Constraint. The two role constraints are the Given A Constraint and the Non-lexical A Constraint.

The One New Argument Constraint states that each clause contains no more than one argument carrying new information, and the One Lexical Argument Constraint indicates that each clause contains no more than one lexical argument. These two quantity constraints are located in parallel dimensions: The One New Argument Constraint is in the pragmatic dimension, while the One Lexical Argument Constraint is in the grammatical dimension. The Given A Constraint claims that the A role typically carries old information, and the Non-lexical A Constraint suggests that lexical arguments do not appear in the A role. These two role constraints are also located in parallel dimensions: The Given A Constraint is in the pragmatic dimension, while the Non-lexical A constraint is in the grammatical dimension. Table 1 summarizes the relation between the constraints and dimensions of preferred argument structure.

The four constraints show the correlation between grammatical roles, pragmatic information, and morphological forms of arguments. These constraints not only hypothesize that there are upper limits to the quantity of information which can be contained in clauses, but also suggest that the distribution of the information shapes an ergative alignment of grammatical roles.

PAS suggests a universal ergative pattern of referent distribution; that is, full lexical noun phrases rarely occur in the A role, but overwhelmingly occur in the S role or the O role. In a pragmatic parallel to this, new information freely appears in the S role or the O role, but not in the A role. In other words, S is aligned with O, as opposed to A. Although Du Bois first derived PAS from narratives in Sakapultek Maya, an ergative language (Du Bois, 1985, 1987), subsequent research has shown that it is a cross-linguistic phenomenon. The patterns of PAS have been documented in a wide array of languages, both of the ergative-absolutive and of the nominative-accusative types. These include languages such as Korean, Nepali, Inuktitut, Finnish, Mapudungun, and Roviana (Clancy, 1993, 2003; Genetti and Crain, 2003; Allen and Schröder, 2003; Helasvuo, 2003; Arnold, 2003; Corston-Oliver, 2003).

However, a number of studies have reported counter-evidence for the hypothesized ergative patterning (O'Dowd, 1990; Chui, 1992; Karkkainen, 1996; Kumagai, 2006; Lin, 2009; Huang, 2012). The data these studies examined were of typologically different languages (e.g., English in O'Dowd, 1990, and Mandarin in Lin, 2009), of different genres (e.g., conversation in Karkkainen, 1996, and narrative in Kumagai, 2006), of different information pressures (e.g., high information pressure in O'Dowd, 1990, and low information pressure in Huang, 2012), and of different types of speakers (e.g., adult speakers in Chui, 1992, and child speakers in Huang, 2012). These studies showed that the ergative patterning was not borne out by the language data they examined. Instead, the data of these studies exhibited an accusative pattern of referent distribution; that is, the S role and the A role were found to contain consistently lower percentages of new mentions and lexical mentions than the O role.

In addition to the ergative patterning, PAS also hypothesizes the existence of universal upper limits of information quantity, as suggested by the One New Argument Constraint and the One Lexical Argument Constraint. That is, there are limits for the number of new mentions and the number of lexical mentions used in a clause in discourse. In contrast to the hypothesized ergative patterning, the hypothesized quantity limits appear to be more robust, supported consistently by evidence across languages (Clancy, 1993, 2003; Genetti and Crain, 2003; Allen and Schröder, 2003; Helasvuo, 2003; Arnold, 2003; Corston-Oliver, 2003). In fact, similar hypotheses concerning information quantity have been proposed by several earlier studies. For example, Pawley and Syder (1983, 2000) proposed the 'one clause at a time' constraint, which suggests that speakers can only give the content of one clause unit at a time. Chafe (1987) proposed the 'one recall at a time' constraint, which suggests that only one inactivated concept can be activated within a single intonation unit. Similarly, Givón (1979) also observed that speakers of Bantu languages restrict one 'bit' of information per proposition or VP. Du Bois' One New Argument Constraint appears to take the matter a step further, suggesting that it is not simply the amount of information, but also the status of the information that is subject to constraint. The constraint appears to have a cognitive basis. As suggested by Du Bois, new information usually is more difficult to process than given information since new information requires a special cognitive effort to bring it into an activated state; because of this, the maximum amount of new information transferred per unit is limited. In addition to the restriction of newness in the One New Argument Constraint, the One Lexical Argument Constraint suggests a restriction on lexicality. In other words, the two quantity constraints indicate two presumably universal upper limits of

information quantity in two parallel dimensions – the limit of information newness in the pragmatic dimension, and the limit of referential lexicality in the grammatical dimension.

1.2. Discourse information measures

While there may be upper limits of information quantity, there appears to be no general constraint specifying the lower limits. Previous studies have also investigated information quantity in terms of the entirety of a given discourse, as in the measure of information pressure (Du Bois, 1987) and the measure of lexical referential density (Bickel, 2003; Stoll and Bickel, 2009). These measures reflect the extent of the informativeness of a given discourse.

Information pressure has to do with the number of new referents introduced into a discourse. When the number of new mentions is larger, the information pressure is higher. Different information pressures also reflect the varying numbers of referents to be tracked in texts of comparable size. As noted by Du Bois (1987), information pressure appears to be related to discourse genres. In some genres, information pressure is often high while in others, pressure is often low. For example, 3rd person stories about strangers may present a higher information pressure than intimate conversation between family members.

While information pressure is related to newness, the measure of lexical referential density concerns lexicality. Parallel to information pressure, lexical referential density has to do with the frequency of the lexical NPs used in a given discourse (Bickel, 2003). Since the introduction of a new entity often motivates the use of a full lexical NP, lexical referential density also appears to be related to discourse genres. In addition, previous studies have shown that lexical referential density is also related to language typology. That is, languages differ in how explicit they are in referencing (Bickel, 2003; Stoll and Bickel, 2009).[1] In other words, languages or genres may differ greatly in lexical referential density.

1.3. Information management: linguistic, pragmatic, and cognitive abilities

The development of the ability to manage information involves linguistic, pragmatic, and cognitive development in children. As seen above, important aspects related to the linguistic domain include the hypothesized universal preferred argument structure (Du Bois, 1987, 2003), typologically different referential density (Bickel, 2003; Stoll and Bickel, 2009), and language-specific referential systems. Children need to acquire these aspects for appropriate referencing.

In the pragmatic domain, previous studies have shown how young children, like adults, choose referential forms according to several discourse-pragmatic factors (Allen, 2000; Clancy, 1993, 1997; Guerriero et al., 2006; Huang, 2011; Narasimhan et al., 2005; Serratrice, 2005). These factors are features of accessibility, such as whether the referents mentioned are present in the physical context of the current conversation, whether the referents have been previously talked about in the conversation, and whether the referents are animate. These studies showed that children tend to use more informative forms (e.g., lexical nouns) for referents that are inaccessible for the features, and less informative forms (e.g., pronouns, null forms) for referents that are accessible for the features, demonstrating children's sensitivity to the pragmatics of reference.

As for the cognitive domain, it has been suggested that managing information in discourse requires an ability by the speaker to take into account the state of mind of the addressee. In other words, children need to develop a 'theory of mind' for appropriate referencing (Gundel, 2009; Gundel and Johnson, 2013; Hughes and Allen, 2013; Rozendaal and Baker, 2010; Salazar Orvig et al., 2010; Skarabela, 2007). 'Theory of mind' refers to the ability to realize that another person is not simply a physical being but a cognitive agent, having his/her own intentions, beliefs, desires, and perspectives that are different from one's own. As suggested by Gundel and Johnson (2013), the appropriate use of referring forms involves the ability to make reasonable assumptions about the addressee's memory and state of attention in relation to the intended referent; i.e., it involves the attribution of mental states to others.

1.4. The present study

In child language research, it is important to analyze not only child speech but also maternal speech and adult speech, in order to understand not only acquisition data, but also input data and target data. It has been widely documented that maternal speech (or child-directed speech in general) differs greatly from adult speech since mothers often adjust their speech when addressing children (e.g., Nelson, 1977; Snow, 1989). While previous studies have characterized maternal speech in terms of syntactic, phonological, or semantic aspects, little has been done to investigate the pragmatic aspects

---

[1] Mandarin allows both subject ellipsis and object ellipsis. Since Mandarin does not have inflection or case markers, the pronominal system is relatively simple, consisting of personal pronouns and demonstrative pronouns. Mandarin NP types include bare nouns, and nouns used with demonstratives, quantifiers, or possessive constructions.

of maternal speech, let alone the aspect of information management. As for child speech, little has been done to study children's skills of information management at the early stages of language acquisition.

This study thus attempts to investigate information management in Mandarin child speech, maternal speech, and adult speech, focusing on three aspects. The first aspect focuses on the patterns of the grammatical alignments in the three types of speech. As shown above, one important hypothesis of PAS is that the subject of a transitive verb (A) is often given and non-lexical, and that the subject of an intransitive verb (S) is aligned with the object of a transitive verb (O) for the accommodation of new mentions and lexical mentions. In other words, PAS suggests an ergative S/O alignment of referent distribution. However, my earlier study of Mandarin child language (Huang, 2012) showed that Mandarin child speech demonstrates an accusative A/S alignment, rather than an S/O alignment, casting doubt on the universality of PAS. In order to better understand the pattern observed in Mandarin child speech, this study aims to investigate whether such grammatical alignment also characterizes Mandarin maternal speech (the input) and adult speech (the target).

The second aspect focuses on the One New Argument Constraint and the One Lexical Argument Constraint. As mentioned earlier, the two quantity constraints indicate the presumably universal upper limits of information quantity in clauses. Huang (2012) showed that the two quantity constraints were observed in Mandarin child speech. This study attempts to further examine whether the constraints also hold for Mandarin maternal speech and adult speech. As mothers usually adjust their speech when addressing children, it is not clear whether the constraints can be observed in both adult-directed speech and child-directed speech. If the constraints hold for Mandarin adult speech, the findings can constitute another piece of cross-linguistic evidence for the universality of the constraints in adult speech. If the constraints hold for Mandarin maternal speech, this may indicate that the constraints are rather robust: Even in adjusted child-directed speech, mothers still follow the constraints of quantity. In other words, this study attempts to examine whether the constraints are supported by data of different speech types. Therefore, the second focus of this study is to examine whether the two constraints can be observed in Mandarin child speech, maternal speech, and adult speech.

The third aspect focuses on comparing Mandarin child speech, maternal speech, and adult speech using two discourse information measures: the measure of information pressure and the measure of lexical referential density. It is speculated that these three types of speech will differ significantly according to these two types of measures, and that the differences will reflect children's developing linguistic ability, mothers' speech adjustments, and adults' mature competence. In other words, the analyses of information pressure and lexical referential density may reveal the characteristics of the three types of speech in information management.

2. Methods

2.1. Participants and data

This study analyzed three types of speech: child speech, maternal speech, and adult speech. The child participants were two Mandarin-speaking girls while the mother participants were the children's mothers, who were native speakers of Mandarin. The two families lived in the northern part of Taiwan. The data collected consisted of eight hours of natural mother–child conversation video-recorded in the participants' homes, with four one-hour sessions with each dyad recorded when the children were at the ages of 2;2, 2;6/2;7, 2;10, and 3;1. The mothers were at the age of 38 and 39 years, respectively at the time of data collection. All of the data were collected in the living rooms of the two homes.

The child data analyzed in this study included the children's speech produced during the total eight hours of data sessions. As for the maternal data, since the mothers produced many more utterances than the children, 20 min of maternal speech from each of the eight sessions, i.e., 160 min in total, were used for analysis. The total numbers of the coded clauses in the child data and in the maternal data (see the coding scheme below) were 2419 and 3351, respectively.

The adult data were adopted from the NCCU Corpus of Spoken Chinese (Chui and Lai, 2008). The adult participants were three dyads of Mandarin-speaking adults each with a close relationship: One of the dyads was composed of a wife (55 years old) and a husband (56 years old); another dyad was composed of a girlfriend (25 years old) and a boyfriend (25 years old), and the other dyad was constituted of two female friends (24 and 25 years old). Each dyad was involved in spontaneous conversation for about 20 min. In other words, the adult data analyzed in the study consisted of 60 min of speech produced by the six adults. The total number of the coded clauses in the adult data was 1688.

2.2. Coding scheme

A coding scheme was adopted to analyze the grammatical alignments and quantity constraints. Each clause which contained an overt finite verb in the data was identified for analysis. The core arguments of each of these verbs, including null arguments, were coded for grammatical roles, information statuses, and referential forms. The coding scheme of this study is as follows:

(1) Grammatical roles
   (a) The A role: It is the subject of a grammatically transitive clause (e.g., the pronoun *wo* 'I' in *wo zai hua meiguihua* 'I am drawing a rose').
   (b) The O role: It is the object of a grammatically transitive clause (e.g., the noun *meiguihua* 'rose' in *wo zai hua meiguihua* 'I am drawing a rose').
   (c) The S role: It is the single argument of a grammatically intransitive clause (e.g., the noun *meimei* 'sister' in *meimei zai ku* 'Sister is crying').
(2) Information statuses
   (a) New: An argument is considered to be new if the referent it denotes has not been mentioned in the preceding 20 utterances (Allen, 2000; Chafe, 1976, 1987; Du Bois, 1987).
   (b) Given: An argument is considered to be given if the referent it denotes has been mentioned in the preceding 20 utterances (Allen, 2000; Chafe, 1976, 1987; Du Bois, 1987).
(3) Referential forms
   (a) Lexical forms: Lexical forms include bare nouns (e.g., *mao* 'cat'), noun phrases (e.g., *hongse de hua* 'red flowers'), and proper names (e.g., *Yiming Shushu* 'Uncle Yiming')
   (b) Non-lexical forms: Non-lexical forms include null forms and pronominal forms (e.g., the pronoun *wo* 'I', the demonstrative *zhe* 'this')

The data were coded by a trained research assistant, who was a native speaker of Mandarin and a graduate student of linguistics. In addition, one fourth of the data were independently coded by another trained research assistant, who was also a native speaker of Mandarin and a graduate student of linguistics. Cohen's Kappa was used to determine the inter-rater reliabilities. The reliabilities for grammatical roles, information statuses, and referential forms were 0.84, 0.83, and 0.92, respectively.

2.3. Discourse information measures

Two types of discourse information measures were included in this study: (1) the measure of information pressure, and (2) the measure of lexical referential density. The details of these two types of measures are presented below.

(1) Information pressure (IP): Information pressure has to do with the number of new referents introduced into a discourse. The value of the information pressure is obtained by calculating the ratio of the new core arguments to the total core arguments in the data (Du Bois, 1987; Durie, 2003), as shown below.

$$\text{Information pressure } (\text{IP}) = \frac{\#\ \text{of new core arguments}}{\#\ \text{of total core arguments}}$$

(2) Lexical referential density (LRD): Lexical referential density has to do with the frequency of the lexical NPs used in a given discourse. The value of the lexical referential density is obtained by calculating the ratio of the lexical core arguments to the total core arguments in the data (Bickel, 2003; Stoll and Bickel, 2009), as shown below.

$$\text{Lexical referential density } (\text{LRD}) = \frac{\#\ \text{of lexical core arguments}}{\#\ \text{of total core arguments}}$$

3. Results

3.1. Grammatical alignments

This section presents the analysis of the grammatical alignments in child speech, maternal speech, and adult speech. More specifically, we analyzed the distributions of new/given information and lexical/non-lexical referential forms in the core arguments of the three types of speech. We attempted to examine whether similar patterns of grammatical alignments can be observed in the three types of speech.

3.1.1. The distributions of new/given information

Table 2a presents the distribution of the new and given mentions within each grammatical role in the child data. The table shows that given mentions occurred more frequently than new mentions in all of the grammatical roles; however, noticeable differences were observed in the distributions of the new and given mentions in the three grammatical roles.

Table 2a
The distribution of new/given information within each grammatical role in the child data.

Child	A		S		O		χ^2
	N	%	N	%	N	%	
New	28	2.16	109	9.71	316	24.38	308.18[***]
Given	1268	97.84	1014	90.29	980	75.62	
Total	1296	100	1123	100	1296	100	

[***] $p < .001$.

Table 2b
Post hoc residual analysis for the child data.

Child	A	S	O
New	−13.7▽	−3.1▽	16.6▲
Given	13.7▲	3.1▲	−16.6▽

Note: ▽, significantly low; ▲, significantly high: $p < .05$ if |adjusted residual| > 1.96; $p < .01$ if |adjusted residual| > 2.58; $p < .001$ if |adjusted residual| > 3.29.

In order to examine whether the differences were statistically significant, a Chi-square analysis was conducted. The result showed that the distributions of the new and given mentions were significantly different across the A, S, and O roles ($\chi^2 = 308.18$, d.f. = 2, $p < .001$), suggesting that the children's use of new and given mentions was influenced by the grammatical roles.

A post hoc residual analysis was further conducted[2] in order to identify the contributions made by the different cells to the resulting Chi-square value. As shown in Table 2b, the figures in all of the cells reached statistical significance, indicating that all of the cells made significant contributions to the result of the Chi-square test. The A role and the S role were significantly less likely to accommodate new mentions, while the O role was significantly more likely to contain new mentions. In contrast, the A role and the S role were significantly more likely to accommodate given mentions, while the O role was significantly less likely to contain given mentions. The results revealed that the S role aligns itself more closely with the A role than it does with the O role in child speech.

Table 3a presents the distribution of the new and given mentions within each grammatical role in the maternal data. As seen in the table, given mentions occurred more frequently than new mentions in all of the grammatical roles; however, the distributions of the new and given mentions appear to present different patterns in the three grammatical roles. A Chi-square analysis was further conducted in order to examine whether the differences were statistically significant. The result showed that the distributions of the new and given mentions were significantly different across the A, S, and O roles ($\chi^2 = 197.05$, d.f. = 2, $p < .001$), suggesting that the mothers' use of new and given mentions was influenced by the grammatical roles.

A post hoc residual analysis was conducted. Table 3b indicates that the figures in all of the cells reached statistical significance, indicating that all of the cells made significant contributions to the result of the Chi-square test. The A role and the S role were significantly less likely to be new, while the O role was significantly more likely to be new. In contrast, the A role and the S role were significantly more likely to be given, while the O role was significantly less likely to be given. The results revealed that the S role aligns itself more closely with the A role than it does with the O role in maternal speech, a pattern similar to that observed in child speech.

As for adult speech, Table 4a shows that while given mentions occurred more frequently than new mentions in all of the grammatical roles, the distributions of the new and given mentions appear to present different patterns in the three grammatical roles. A Chi-square analysis was also conducted, and the result showed that the distributions of the new and given mentions were significantly different across the three grammatical roles ($\chi^2 = 218.51$, d.f. = 2, $p < .001$), suggesting that the adults' use of new and given mentions was influenced by the grammatical roles.

A post hoc residual analysis was further conducted. As shown in Table 4b, the figures in all of the cells reached statistical significance, indicating that all of the cells made significant contributions to the result of the Chi-square test. The A role and the S role were significantly less likely to be new, while the O role was significantly more likely to be new. In contrast, the A role and the S role were significantly more likely to be given, while the O role was significantly less likely to

---

[2] A residual is the difference between the observed and expected values for a cell. A residual analysis identifies those specific cells making the greatest contribution to the Chi-square test result.

Table 3a
The distribution of new/given information within each grammatical role in the maternal data.

Maternal	A		S		O		χ^2
	N	%	N	%	N	%	
New	20	1.31	91	4.99	206	13.47	197.05[***]
Given	1509	98.69	1731	95.01	1323	86.53	
Total	1529	100	1822	100	1529	100	

[***] $p < .001$.

Table 3b
Post hoc residual analysis for the maternal data.

Maternal	A	S	O
New	−9.9▽	−3.3▽	13.4▲
Given	9.9▲	3.3▲	−13.4▽

Note: ▽, significantly low; ▲, significantly high: $p < .05$ if |adjusted residual| > 1.96; $p < .01$ if |adjusted residual| > 2.58; $p < .001$ if |adjusted residual| > 3.29.

Table 4a
The distribution of new/given information within each grammatical role in the adult data.

Adult	A		S		O		χ^2
	N	%	N	%	N	%	
New	22	3.02	132	13.75	225	30.91	218.51[***]
Given	706	96.98	828	86.25	503	69.09	
Total	728	100	960	100	728	100	

[***] $p < .001$.

Table 4b
Post hoc residual analysis for the adult data.

Adult	A	S	O
New	−11.2▽	−2.1▽	13.5▲
Given	11.2▲	2.1▲	−13.5▽

Note: ▽, significantly low; ▲, significantly high: $p < .05$ if |adjusted residual| > 1.96; $p < .01$ if |adjusted residual| > 2.58; $p < .001$ if |adjusted residual| > 3.29.

be given. The results revealed that the S role aligns itself more closely with the A role than it does with the O role in adult speech, a pattern similar to those observed in child speech and in maternal speech.

In sum, an A/S alignment was observed in the distributions of new and given information in Mandarin child speech, maternal speech, and adult speech. The pattern was inconsistent with the S/O alignment suggested by PAS.

3.1.2. The distributions of referential forms

The referential forms used in the child data, the maternal data, and the adult data were also analyzed. Table 5a presents the distribution of the lexical and non-lexical mentions within each grammatical role in the child data. The table shows that in the child data, non-lexical forms were used more frequently than lexical forms in all of the grammatical roles; however, noticeable differences were observed in the distributions of the lexical and non-lexical mentions in the three grammatical roles. In order to examine whether the differences were statistically significant, a Chi-square analysis was conducted. The result showed that the distributions of the lexical and non-lexical mentions were significantly different across the A, S, and O roles ($\chi^2 = 514.56$, d.f. = 2, $p < .001$), suggesting that the children's use of referential forms was influenced by the grammatical roles.

A post hoc residual analysis was further conducted in order to identify the contributions made by the different cells to the resulting Chi-square value. As shown in Table 5b, the figures in all of the cells reached statistical significance,

Table 5a
The distribution of lexical/non-lexical forms within each grammatical role in the child data.

Child	A		S		O		χ^2
	N	%	N	%	N	%	
Lexical	94	7.25	156	13.89	542	41.82	514.56***
Non-lexical	1202	92.75	967	86.11	754	58.18	
Total	1296	100	1123	100	1296	100	

*** $p < .001$.

Table 5b
Post hoc residual analysis for the child data.

Child	A	S	O
Lexical	−15.3▽	−7.3▽	22.3▲
Non-lexical	15.3▲	7.3▲	−22.3▽

Note: ▽, significantly low; ▲, significantly high; $p < .05$ if |adjusted residual| > 1.96; $p < .01$ if |adjusted residual| > 2.58; $p < .001$ if |adjusted residual| > 3.29.

indicating that all of the cells made significant contributions to the result of the Chi-square test. The A role and the S role were significantly less likely to accommodate lexical mentions, while the O role was significantly more likely to contain lexical mentions. In contrast, the A role and the S role were significantly more likely to accommodate non-lexical mentions, while the O role was significantly less likely to contain non-lexical mentions. The results revealed that the S role aligns itself more closely with the A role than it does with the O role in child speech.

Table 6a presents the distribution of the lexical and non-lexical mentions within each grammatical role in the maternal data. As seen in the table, non-lexical mentions occurred more frequently than lexical mentions in all of the grammatical roles; however, the distributions of the lexical and non-lexical mentions appear to present different patterns in the three grammatical roles. A Chi-square analysis was further conducted in order to examine whether the differences were statistically significant. The result showed that the distributions of the lexical and non-lexical mentions were significantly different across the A, S, and O roles ($\chi^2 = 412.42$, d.f. = 2, $p < .001$), suggesting that the mothers' use of lexical and non-lexical mentions was influenced by the grammatical roles.

A post hoc residual analysis was conducted. Table 6b shows that the figures in all of the cells reached statistical significance, indicating that all of the cells made significant contributions to the result of the Chi-square test. The A role and the S role were significantly less likely to be lexical, while the O role was significantly more likely to be lexical. In contrast,

Table 6a
The distribution of lexical/non-lexical forms within each grammatical role in the maternal data.

Maternal	A		S		O		χ^2
	N	%	N	%	N	%	
Lexical	238	15.57	326	17.89	673	44.02	412.42***
Non-lexical	1291	84.43	1496	82.11	856	55.98	
Total	1529	100	1822	100	1529	100	

*** $p < .001$.

Table 6b
Post hoc residual analysis for the maternal data.

Maternal	A	S	O
Lexical	−10.6▽	−9.2▽	20.2▲
Non-lexical	10.6▲	9.2▲	−20.2▽

Note: ▽, significantly low; ▲, significantly high; $p < .05$ if |adjusted residual| > 1.96; $p < .01$ if |adjusted residual| > 2.58; $p < .001$ if |adjusted residual| > 3.29.

Table 7a
The distribution of lexical/non-lexical forms within each grammatical role in the adult data.

Adult	A		S		O		χ^2
	N	%	N	%	N	%	
Lexical	70	9.62	325	33.85	431	59.20	397.88[***]
Non-lexical	658	90.38	635	66.15	297	40.80	
Total	728	100	960	100	728	100	

*** $p < .001$.

Table 7b
Post hoc residual analysis for the adult data.

Adult	A	S	O
Lexical	−16.7▽	−0.3 n.s.	17▲
Non-lexical	16.7▲	0.3 n.s.	−17▽

Note: ▽, significantly low; ▲, significantly high; $p < .05$ if |adjusted residual| > 1.96; $p < .01$ if |adjusted residual| > 2.58; $p < .001$ if |adjusted residual| > 3.29.

the A role and the S role were significantly more likely to be non-lexical, while the O role was significantly less likely to be non-lexical. The results revealed that the S role aligns itself more closely with the A role than it does with the O role in maternal speech, a pattern similar to that observed in child speech.

As for adult speech, Table 7a shows that while non-lexical mentions occurred more frequently than lexical mentions in all of the grammatical roles, the distributions of the lexical and non-lexical mentions appear to present different patterns in the three grammatical roles. A Chi-square analysis was also conducted, and the result showed that the distributions of the lexical and non-lexical mentions were significantly different across the three grammatical roles (χ^2 = 397.88, d.f. = 2, $p < .001$), suggesting that the adults' use of lexical and non-lexical mentions was influenced by the grammatical roles.

A post hoc residual analysis was further conducted. Table 7b shows that it was mainly the A role and the S role that contributed to the resulting value of the Chi-square test. As shown in the table, the A role was significantly less likely to be lexical, while the O role was significantly more likely to be lexical. In contrast, the A role was significantly more likely to be non-lexical, while the O role was significantly less likely to be non-lexical. The distribution in the S role, however, did not reach significance.

In sum, an A/S alignment was observed in the distributions of the lexical and non-lexical information in Mandarin child speech and maternal speech. Adult speech, however, did not demonstrate a significant pattern of alignment.

This section has analyzed the distributions of information statuses and referential forms in child, maternal, and adult speech. The results showed that an ergative pattern was not observed in the three types of speech, casting doubt on the universality of PAS. Instead, Mandarin-speaking children, mothers, and adults tend to align the S role with the A role, rather than with the O role, demonstrating an accusative alignment in their speech.

Interestingly, although the three types of speech demonstrated similar patterns in grammatical alignment, some noticeable differences were also observed in the distributions. For example, it appears that the proportion of given referents is higher for maternal speech than for either child or adult speech, as can be seen in Tables 2a, 3a and 4a. In addition, it appears that the proportion of non-lexical mentions is considerably lower for adult speech than for either child or maternal speech, as can be seen in Tables 5a, 6a and 7a. Further analysis concerning the differences between the three types of speech is presented in Section 3.3.

3.2. Quantity constraints

This section presents the analysis of information quantity in child speech, maternal speech, and adult speech. More specifically, we tested the One New Argument Constraint and the One Lexical Argument Constraint proposed by Du Bois (1987) by analyzing the numbers of new/given mentions and the numbers of lexical/non-lexical mentions in the clauses of the three types of data. These constraints indicate the presumably universal upper limits of information quantity contained in clauses. We attempted to examine whether such constraints can be observed in Mandarin child speech, maternal speech, and adult speech.

Table 8
Frequency of clauses with zero, one, or two new arguments.

New	Child		Maternal		Adult	
	N	%	N	%	N	%
0	1982	81.93	3040	90.72	1321	78.26
1	421	17.40	305	9.10	355	21.03
2	16	0.66	6	0.18	12	0.71
Total	2419	100	3351	100	1688	100

Table 9
Transitivity and frequency of clauses with zero, one, or two new arguments.

New	Child				Maternal				Adult			
	Transitive		Intransitive		Transitive		Intransitive		Transitive		Intransitive	
	N	%	N	%	N	%	N	%	N	%	N	%
0	968	74.69	1014	90.29	1309	85.61	1731	95.01	493	67.72	828	86.25
1	312	24.07	109	9.71	214	14.00	91	4.99	223	30.63	132	13.75
2	16	1.23	0	0.00	6	0.39	0	0.00	12	1.65	0	0.00
Total	1296	100	1123	100	1529	100	1822	100	728	100	960	100

3.2.1. One New Argument Constraint

The One New Argument Constraint indicates that each clause contains no more than one argument carrying new information. Table 8 displays the frequency of the clauses with zero, one, or two new arguments in the data. As seen in the table, the percentages of the clauses with two new arguments in the three types of data were very low. They were 0.66%, 0.18%, and 0.71% in the child data, the maternal data, and the adult data, respectively. The majority of the clauses in the data contained no new arguments at all (81.93%, 90.72%, and 78.26%). The results thus confirmed the One New Argument Constraint. In other words, the children, the mothers, and the adults all tended to produce clauses which contained no more than one new argument.

Since only transitive clauses can have more than one core argument, further analysis was conducted to analyze the clauses in the data in terms of transitivity. As seen in Table 9, only 1.23%, 0.39%, and 1.65% of the transitive clauses in the three types of speech had two new arguments. Most of the transitive clauses in the data had either one or zero new argument, and the percentages of the clauses with zero new argument were much higher than those of the clauses with one new argument in the three types of speech. As for the intransitive clauses, only a few intransitive clauses had one new argument; the overwhelming majority of the intransitive clauses had zero new argument (90.29%, 95.01%, and 86.25%). The results thus confirmed the One New Argument Constraint. In other words, the children, the mothers, and the adults all tended to produce clauses which contained no more than one new argument.

3.2.2. One Lexical Argument Constraint

The One Lexical Argument Constraint claims that each clause contains no more than one lexical argument. Table 10 presents the frequency of the clauses with zero, one, or two lexical arguments in the data. As seen in the table, only 1.65%, 3.04%, and 2.61% of the clauses in the three types of speech contained two lexical arguments. The majority of the

Table 10
Frequency of clauses with zero, one, or two lexical arguments.

Lexical	Child		Maternal		Adult	
	N	%	N	%	N	%
0	1667	68.91	2216	66.13	906	53.67
1	712	29.43	1033	30.83	738	43.72
2	40	1.65	102	3.04	44	2.61
Total	2419	100	3351	100	1688	100

Table 11
Transitivity and frequency of clauses with zero, one, or two lexical arguments.

| Lexical | Child | | | | Maternal | | | | Adult | | | |
| | Transitive | | Intransitive | | Transitive | | Intransitive | | Transitive | | Intransitive | |
	N	%	N	%	N	%	N	%	N	%	N	%
0	700	54.01	967	86.11	720	47.09	1496	82.11	271	37.23	635	66.15
1	556	42.90	156	13.89	707	46.24	326	17.89	413	56.73	325	33.85
2	40	3.09	0	0.00	102	6.67	0	0.00	44	6.04	0	0.00
Total	1296	100	1123	100	1529	100	1822	100	728	100	960	100

clauses in the data had no lexical arguments at all (68.91%, 66.13%, and 53.67%). The results thus supported the One Lexical Argument Constraint. In other words, the children, the mothers, and the adults all tended to produce clauses which contained no more than one lexical argument.

Further analysis was conducted to analyze the clauses in the data in terms of transitivity. The results are shown in Table 11. As seen in the table, only 3.09%, 6.67%, and 6.04% of the transitive clauses in the three types of speech had two lexical arguments. Most of the transitive clauses in the data had either one or zero lexical argument. As for the intransitive clauses, while a few of the intransitive clauses had one lexical argument, the vast majority of the intransitive clauses had zero lexical argument (86.11%, 82.11%, and 66.15%). The results thus supported the One Lexical Argument Constraint. In other words, the children, the mothers, and the adults all tended to produce clauses which contained no more than one lexical argument.

In short, the two constraints were supported by the results of the present study. It was notably rare that the Mandarin-speaking children, mothers, or adults produced clauses with more than one new argument or with more than one lexical argument. In other words, these constraints apply not only to adult speech, as suggested by previous studies, but also to maternal speech and to child speech. The upper limits of information quantity can be observed in all of the three types of speech.

While all of the three types of speech followed the One New Argument Constraint and the One Lexical Argument Constraint, it is worth noting that the distributions presented in the tables above also reveal some differences between the three types of speech. For example, Tables 8 and 9 show that maternal speech appears to have a higher proportion of clauses with zero new argument compared to the other two types of speech. In addition, Tables 10 and 11 show that adult speech appears to have fewer clauses with zero lexical arguments than the other two types of speech. Further analysis concerning the differences between the three types of speech is presented in Section 3.3.

3.3. Discourse information measures

As seen above, while child speech, maternal speech, and adult speech demonstrated similar patterns in the analyses of grammatical alignments and quantity constraints, a closer look at the three types of speech revealed some noticeable differences in the proportions of new information and lexical information. In order to obtain a more complete picture, this section presents the analysis of two types of discourse information measures. More specifically, we measured the values of information pressure (Du Bois, 1987; Durie, 2003) and the values of lexical referential density (Bickel, 2003; Stoll and Bickel, 2009) for the three types of speech.

3.3.1. Information pressure (IP) and lexical referential density (LRD)

The IP values for child, maternal, and adult speech are shown in Table 12. The three values were significantly different ($\chi^2 = 164.62$, d.f. = 2, $p < .001$). A post hoc multiple comparison test (Marascuilo and McSweeney, 1977) showed that the

Table 12
Information pressure (IP) in child speech, maternal speech and adult speech.

	Child	Maternal	Adult
IP (new/total)	0.12 (453/3715)	0.06 (317/4880)	0.16 (379/2416)
χ^2		164.62[***]	
Post hoc		M < C, C < A, M < A (M < C < A)	

[***] $p < .001$.

Table 13
Lexical referential density (LRD) in child speech, maternal speech and adult speech.

	Child	Maternal	Adult
LRD (lexical/total)	0.21 (792/3715)	0.25 (1237/4880)	0.34 (826/2416)
χ^2		127.79***	
Post hoc		C < M, C < A, M < A (C < M < A)	

*** $p < .001$.

Table 14
Summary table of IP and LRD.

	Child	Maternal	Adult
IP	Middle	Lowest	Highest
LRD	Lowest	Middle	Highest

IP value for maternal speech was significantly lower than that for either child or adult speech, and that the IP value for child speech was significantly lower than that for adult speech. In other words, adult speech contained the largest proportion of new mentions, while maternal speech contained the smallest proportion of new mentions.

The LRD values for child, maternal, and adult speech are shown in Table 13. The three values were also significantly different (χ^2 = 127.79, d.f. = 2, $p < .001$). A post hoc multiple comparison test (Marascuilo and McSweeney, 1977) showed that the LRD value for child speech was significantly lower than that for either maternal or adult speech, and that the IP value for maternal speech was significantly lower than that for adult speech. In other words, adult speech contained the largest proportion of lexical mentions, while child speech contained the smallest proportion of lexical mentions.

As seen in Tables 12 and 13, adult speech had both the highest IP and the highest LRD when compared with the other two types of speech. In other words, adult speech provided new mentions and lexical forms more frequently than maternal speech or child speech. Interestingly, the results showed that child speech had a higher IP than maternal speech, but that maternal speech had a higher LRD than child speech. In other words, the children were more likely than their mothers to provide new information, while the mothers were more likely than the children to supply lexical forms.

To summarize, the IPs and the LRDs for the three types of speech, which indicate the proportions of new mentions and lexical mentions in the data, can be demonstrated in Table 14. Adult speech, not surprisingly, was characterized by the highest IP and the highest LRD; maternal speech was characterized by the lowest IP and a middle LRD, and child speech, on the other hand, was characterized by a middle IP and the lowest LRD.

3.3.2. New mentions and lexical mentions

To further investigate how the distributions of new mentions and lexical mentions differed in the three types of speech, analysis was conducted to obtain the ratios of new mentions to lexical mentions in the three types of speech, as shown in Table 15. The results showed that there were more lexical mentions than new mentions in all of the three types of speech. However, the difference between the number of lexical mentions and the number of new mentions was larger in maternal speech than in the other two types of speech. The mothers produced about four times as many lexical mentions as they did new mentions; the adults produced more than twice as many lexical mentions as they did new mentions, and the children produced less than twice as many lexical mentions as they did new mentions.

While it has often been suggested that the association between new information and lexical mentions is strong (Du Bois, 1987; Clancy, 2003), Table 15 shows that this association is only partial, and that the strength of the association may vary according to the type of speech. Since the mothers produced considerably more lexical mentions than new mentions, we suspect that it may be more likely for the mothers than for the children or the adults to use lexical forms for referents which were given. Further analysis was conducted to investigate the proportions of lexical mentions used

Table 15
Ratio of lexical mentions to new mentions in child speech, maternal speech, and adult speech.

	Child	Maternal	Adult
New mentions	453	317	379
Lexical mentions	792	1237	826
Ratio	1:1.75	1:3.90	1:2.18

Table 16
The proportions of lexical mentions used for new referents and for given referents.

Lexical	Child		Maternal		Adult		χ^2	Post hoc
	N	%	N	%	N	%		
New	288	36.36	231	18.67	366	44.31	166.93***	M < C, C < A, M < A (M < C < A)
Given	504	63.64	1006	81.33	460	55.69		C < M, A < C, A < M (A < C < M)
Total	792	100	1237	100	826	100		

*** $p < .001$.

Table 17
The proportions of new referents mentioned with lexical forms and with non-lexical forms.

New	Child		Maternal		Adult		χ^2	Post hoc
	N	%	N	%	N	%		
Lexical	288	63.58	231	72.87	366	96.57	131.20***	C < M, C < A, M < A (C < M < A)
Non-lexical	165	36.42	86	27.13	13	3.43		M < C, A < C, A < M (A < M < C)
Total	453	100	317	100	379	100		

*** $p < .001$.

for referring to new referents and for referring to given referents in the three types of speech in order to examine whether maternal speech was characterized by a higher frequency of given lexical mentions. The results are shown in Table 16.

As shown in Table 16, the distributions in the three types of speech were significantly different (χ^2 = 166.93, d.f. = 2, $p < .001$). A post hoc multiple comparison test (Marascuilo and McSweeney, 1977) showed that the percentage of given lexical mentions in adult speech was significantly lower than that in either child or maternal speech, and that the percentage in child speech was significantly lower than that in maternal speech. In other words, maternal speech contained the largest proportion of given lexical mentions, while adult speech contained the smallest proportion of given lexical mentions. It appears that maternal speech was characterized by a higher frequency of given lexical mentions compared with the other two types of speech.

Excerpt 1 demonstrates how a lexical form was used to refer to a given referent in maternal speech. As seen in this example, the mother used the lexical form 'Barbie' several times to refer to a doll, which was a given referent in the context.

Excerpt 1[3]

```
*MOT:  a       nide Babi    lei? ←
       PRT   your Barbie  PRT
       '(Where's) your Barbie?''
*CHI:  en.
       'Um.'
*MOT:  Babi     tang zai dishang  shi-bu-shi? ←
       Barbie  lie    on  floor      right
       'Barbie is lying on the floor, right?'
*CHI:  en.
       'Um.'
*MOT:  ba   Babi    jian  qilai. ←
       BA  Barbie  pick  up
       'Pick Barbie up.'
```

In addition to the analysis of given lexical mentions, a related analysis is to examine the use of non-lexical new mentions in the three types of speech. Analysis was conducted to examine the proportions of new referents which were mentioned with lexical forms and with non-lexical forms in the three types of speech. The results are shown in Table 17. As seen in Table 17, the distributions of the lexical forms and non-lexical forms in the three types of speech were significantly

---

[3] Gloss abbreviations: BA/ba; PRT/Particle.

different (χ^2 = 131.20, d.f. = 2, p < .001). A post hoc multiple comparison test (Marascuilo and McSweeney, 1977) showed that the percentage of non-lexical new mentions in adult speech was significantly lower than that in either maternal or child speech, and that the percentage in maternal speech was significantly lower than that in child speech. In other words, child speech contained the largest proportion of non-lexical new mentions, while adult speech contained the smallest proportion of non-lexical new mentions. It appears that child speech was characterized by a higher frequency of non-lexical new mentions compared with the other two types of speech.

Excerpt 2 demonstrates how a new referent was mentioned in child speech with a non-lexical form, a null form in this case. As seen in this example, the child used a null form when introducing a new referent. The mother responded with a question for clarification.

Excerpt 2

```
*CHI:   mama    wo  yao   chi  Ø. ←
        Mommy  I    want  eat  Ø
        'Mommy, I want to eat Ø.'
*MOT:   chi   sheme?
        eat   what
        'What do you want to eat?'
```

In sum, the analyses showed that maternal speech was characterized by a frequent use of given lexical mentions; child speech was characterized by a frequent use of non-lexical new mentions. Adult speech, on the other hand, demonstrated a strong association between new referents and lexical forms.

4. Discussion and conclusion

This study has presented an analysis of information management in Mandarin child speech, maternal speech, and adult speech. Similarities and differences were found between the three types of speech in relation to grammatical alignments, quantity constraints, and discourse information measures.

The results showed that rather than demonstrating an S/O alignment, as predicted by PAS, the three types of speech demonstrated an A/S alignment, revealing that Mandarin is characterized by an accusative pattern, and also that children are sensitive to this pattern early on. The results suggested that an ergative pattern of grammatical alignment may not be linguistically universal.

As for the quantity constraints, this study showed that the One New Argument Constraint and the One Lexical Argument Constraint are supported by Mandarin child speech, maternal speech, and adult speech. While previous studies have shown that the quantity constraints are supported by adult speech cross-linguistically, this study provided evidence that these constraints can also be observed in child speech at the early stages of language acquisition. In addition, while maternal speech was characterized by a larger proportion of lexical mentions than child speech and adult speech, the results demonstrated that the mothers also follow the upper limits of lexical quantity indicated by the constraints.

While the three types of speech showed similar patterns in grammatical alignments and quantity constraints, this study further demonstrated how measures of discourse information can reveal the differences between the three types of speech in information management, and how these differences may reflect the characteristics of the three types of speech. As seen in the results, adult speech was characterized by the highest IP and the highest LRD; it was also characterized by a strong association between new referents and lexical forms. The results appear to reflect the nature of adult–adult conversation: Both the speaker and the listener in adult–adult conversation are competent language users. Conversational topics in adult–adult conversation, unlike those in conversation involving children, usually go beyond the 'here-and-now'. New referents and lexical expressions thus tend to be used more often in adult speech in order to meet the communicative needs of adult speakers.

Maternal speech is directed to linguistically and cognitively developing children. While mothers have mature linguistic ability, they tend to adjust their speech when interacting with children. As seen in the results, maternal speech was characterized by the lowest IP. It appears that the mothers tended to restrict the introduction of new referents in order to ease the information processing load for the children. Maternal speech was also characterized by a large proportion of given lexical mentions; that is, the mothers tended to use informative forms (i.e., lexical forms) to refer to given referents. This may result from the mothers' attempt to ensure the children's comprehension of the referents mentioned. The results appear to reflect the mothers' attempt to facilitate their communication with the children. This finding is consistent with those reported in Hughes and Allen (2015) and in Clark and Bernicot (2008). Hughes and Allen found a high proportion of

lexical NPs for accessible referents in their adult-child data, and they showed that the phenomenon was largely due to the repetition of a part of the previous utterance. Clark and Bernicot also found a tendency for adults to repeat lexical items in child-directed speech. They showed that adults repeat lexical items to check and confirm children's communicative intentions, to correct errors in what children have said, or to signal that they are attending to children's utterances. In addition, it appears that the high proportions of given lexical mentions in maternal speech are also related to what Küntay and Slobin (1996, 2002) have called ''variation sets''. Variation sets are a certain kind of discourse form in child-directed speech, characterized by a sequence of utterances with a constant intention, but varying forms (See Excerpt 1 for example). The interactive function of variation sets, as suggested by Küntay and Slobin, is to maximize the chance of comprehension and/or compliance on the part of the child through the repetition of the same content. Variation sets thus appear to motivate the frequent use of lexical mentions for given information. Interestingly, the characteristic of a large proportion of given lexical mentions in maternal speech was also found in teacher talk, as shown in the results of Kumpf (2003). It appears that this characteristic reflects the similarities between maternal speech and teacher talk: Both mothers and teachers speak to addressees with less competence, and they both share the goal of trying to ensure the comprehension of their addressees.

As for child speech, it was characterized by the lowest LRD. That is, the children tended to use pronominal forms or even null forms, which are less informative than lexical forms. Child speech was also characterized by a large proportion of non-lexical new mentions. In other words, when the children introduced new referents, the referential forms they used may be non-informative. One possible explanation for this finding is that the children were pragmatically immature and not sensitive to the perspective of their interlocutors. However, Skarabela et al. (2013) reported that this phenomenon may be associated with children's sensitivity to the presence or absence of joint attention. They showed that children use pronominal and null forms for new referents largely when those referents are already clarified through the presence of joint attention. As also suggested by Allen et al. (2015), children are sensitive to subtle aspects of information flow, and they attend to the interplay between accessibility factors (e.g., the interplay between 'newness' and 'joint attention') when choosing referential forms. In other words, the large proportion of non-lexical new mentions in child speech may not result from children's pragmatic deficiency; instead, most of these non-lexical new mentions may show children's adjustment of their messages for the interlocutors, thus reflecting their developing ability to acquire a theory of mind. Further studies are needed in order to better understand how the interplay between accessibility factors influences children's referential choice.

This study has shown some similarities and differences in information management between Mandarin child speech, maternal speech, and adult speech. The three types of speech were similar in that they demonstrated an accusative pattern of grammatical alignment, and that they followed the quantity constraints. However, they were different in discourse information measures, i.e., in the value of information pressure and in the value of lexical referential density. By examining the three types of speech, this study has shown how information management is associated with children's developing linguistic ability, mothers' speech adjustments, and adults' mature linguistic competence.

Acknowledgements

References

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: Du Bois, J.W., et al. (Eds.), Preferred Argument Structure: Grammar as Architecture for Function. John Benjamins, Amsterdam, pp. 301–338.

Allen, S.E.M., et al., 2015. The role of cognitive accessibility in children's referential choice. In: Serratrice, L., Allen, S.E.M. (Eds.), The Acquisition of Reference. Benjamins, Amsterdam, pp. 123–153.

Arnold, J.E., 2003. Multiple constraints on reference forms: null, pronominal, and full reference in Mapudungun. In: Du Bois, J.W., et al. (Eds.), Preferred Argument Structure: Grammar as Architecture for Function. John Benjamins, Amsterdam, pp. 225–246.

Bickel, B., 2003. Referential density in discourse and syntactic typology. Language 79, 708–736.

Chafe, W., 1976. Givenness, contrastiveness, definiteness, subjects, topics, and point of view. In: Li, C. (Ed.), Subject and Topic. Academic Press, London, pp. 25–55.

Chafe, W., 1987. Cognitive constraints on information flow. In: Tomlin, R.S. (Ed.), Coherence and Grounding in Discourse. Benjamins, Amsterdam, pp. 21–51.

Chui, K., 1992. Preferred argument structure for discourse understanding. In: Proceedings of the 14th conference on Computational Linguistics (COLING92), vol. 4, Association for Computational Linguistics, pp. 1142–1146.

Chui, K., Lai, H.-l., 2008. The NCCU corpus of spoken Chinese: Mandarin, Hakka, and Southern Min. Taiwan J. Linguist. 6 (2), 119–144.

Clancy, P., 1993. Preferred argument structure in Korean acquisition. In: Clark, E.V. (Ed.), The Proceedings of the Twenty-Fifth Child Language Research Forum. CSLI, Stanford, CA, pp. 307–314.

Clancy, P., 1997. Discourse motivations for referential choice in Korean acquisition. In: Sohn, H., Haig, J. (Eds.), Japanese/Korean Linguistics, vol. 6. CSLI, Stanford, CA, pp. 639–657.

Clancy, P., 2003. The lexicon in interaction: developmental origins of preferred argument structure in Korean. In: Du Bois, J.W., et al. (Eds.), Preferred Argument Structure: Grammar as Architecture for Function. John Benjamins, Amsterdam, pp. 81–108.

Clark, E.V., Bernicot, J., 2008. Repetition as ratification: how parents and children place information in common ground. J. Child Lang. 35, 349–371.

Corston-Oliver, S.H., 2003. Core arguments and the inversion of the nominal hierarchy in Roviana. In: Du Bois, J.W., et al. (Eds.), Preferred Argument Structure: Grammar as Architecture for Function. John Benjamins, Amsterdam, pp. 273–300.

Du Bois, J.W., 1985. Competing motivations. In: Haiman, J. (Ed.), Iconicity in Syntax. John Benjamins, Amsterdam, pp. 343–365.

Du Bois, J.W., 1987. The discourse basis of ergativity. Language 63, 805–855.

Du Bois, J.W., 2003. Argument structure: grammar in use. In: Du Bois, J.W., et al. (Eds.), Preferred Argument Structure: Grammar as Architecture for Function. John Benjamins, Amsterdam, pp. 11–60.

Du Bois, J.W., et al., 2003. Preferred Argument Structure: Grammar as Architecture for Function. John Benjamins, Amsterdam.

Durie, M., 2003. New light on information pressure: information conduits, ''escape valves'', and role alignment stretching. In: Du Bois, J.W., et al. (Eds.), Preferred Argument Structure: Grammar as Architecture for Function. John Benjamins, Amsterdam, pp. 159–196.

Genetti, C., Crain, L.D., 2003. Beyond preferred argument structure: sentences, pronouns, and given referents in Nepali. In: Du Bois, J.W., et al. (Eds.), Preferred Argument Structure: Grammar as Architecture for Function. John Benjamins, Amsterdam, pp. 197–224.

Givón, T., 1979. On Understanding Grammar. Academic Press, New York.

Guerriero, S., et al., 2006. The development of referential choice in English and Japanese: a discourse-pragmatic perspective. J. Child Lang. 33, 823–857.

Gundel, J.K., 2009. Children's use of referring expressions: what can it tell us about theory of mind? Cognit. Crit. 1, 73–100.

Gundel, J.K., Johnson, K., 2013. Children's use of referring expressions in spontaneous discourse: implications for theory of mind development. J. Pragmat. 56, 43–57.

Helasvuo, M.-L., 2003. Argument splits in Finnish grammar and discourse. In: Du Bois, J.W., et al. (Eds.), Preferred Argument Structure: Grammar as Architecture for Function. John Benjamins, Amsterdam, pp. 247–272.

Huang, C.-c., 2011. Referential choice in Mandarin child language: a discourse-pragmatic perspective. J. Pragmat. 43, 2057–2080.

Huang, C.-c., 2012. Preferred argument structure in Mandarin child language. Taiwan J. Linguist. 10 (2), 119–168.

Hughes, M.E., Allen, S.E.M., 2013. The effect of individual discourse-pragmatic features on referential choice in child English. J. Pragmat. 56, 15–30.

Hughes, M.E., Allen, S.E.M., 2015. The incremental effect of discourse-pragmatic sensitivity on referential choice in the acquisition of a first language. Lingua 155, 43–61.

Karkkainen, E., 1996. Preferred argument structure and subject role in American English conversational discourse. J. Pragmat. 25, 675–701.

Kumagai, Y., 2006. Information management in intransitive subjects: some implications for the preferred argument structure theory. J. Pragmat. 38, 670–694.

Kumpf, L.E., 2003. Genre and preferred argument structure: sources of argument structure in classroom discourse. In: Du Bois, J.W., et al. (Eds.), Preferred Argument Structure: Grammar as Architecture for Function. John Benjamins, Amsterdam, pp. 109–130.

Küntay, A., Slobin, D.I., 1996. Listening to a Turkish mother: some puzzles for acquisition. In: Slobin, D.I., et al. (Eds.), Social Interaction, Social Context, and Language: Essays in Honor of Susan Ervin-Tripp. Lawrence Erlbaum Associates, Hillsdale, NJ, pp. 265–286.

Küntay, A., Slobin, D.I., 2002. Putting interaction back into child language: examples from Turkish. Psychol. Lang. Commun. 6, 5–14.

Lin, W.-h., 2009. Preferred argument structure in Chinese: a comparison among conversations, narratives and written texts. In: Xiao, Y. (Ed.), The Proceedings of the 21st North American Conference on Chinese Linguistics (NACCL-21), vol. 2. pp. 341–357.

Marascuilo, L.A., McSweeney, M., 1977. Nonparametric and Distribution-Free Methods for the Social Science. Brooks/Cole, Monterey, CA.

Narasimhan, B., et al., 2005. Argument realization in Hindi caregiver-child discourse. J. Pragmat. 37, 461–495.

Nelson, K., 1977. Facilitating children's acquisition of syntax. Dev. Psychol. 13, 101–107.

O'Dowd, E., 1990. Discourse pressure, genre and grammatical alignment – after Du Bois. Stud. Lang. 14 (2), 365–403.

Pawley, A., Syder, F.H., 1983. Natural selection in syntax: notes on adaptive variation and change in vernacular and literary grammar. J. Pragmat. 7, 551–579.

Pawley, A., Syder, F.H., 2000. The one-clause-at-a-time hypothesis. In: Riggenbach, H. (Ed.), Perspectives on Fluency. University of Michigan Press, Ann Arbor, MI, pp. 163–199.

Rozendaal, M.I., Baker, A.E., 2010. The acquisition of reference: pragmatic aspects and the influence of language input. J. Pragmat. 42, 1866–1879.

Salazar Orvig, A., et al., 2010. Dialogical beginnings of anaphora: the use of third person pronouns before the age of 3. J. Pragmat. 42, 1842–1865.

Serratrice, L., 2005. The role of discourse pragmatics in the acquisition of subjects in Italian. Appl. Psycholinguist. 26, 437–462.

Skarabela, B., 2007. Signs of early social cognition in children's syntax: the case of joint attention in argument realization in child Inuktitut. Lingua 117, 1837–1857.

Skarabela, B., et al., 2013. Joint attention helps explain why children omit new arguments. J. Pragmat. 56, 5–14.

Snow, C.E., 1989. Understanding social interaction and language acquisition: sentences are not enough. In: Bornstein, M.H., Bruner, J.S. (Eds.), Interaction in Human Development. Lawrence Erlbaum, Hillsdale, NJ, pp. 83–104.

Stoll, S., Bickel, B., 2009. How deep are differences in referential density? In: Lieven, E., et al. (Eds.), Crosslinguistic Approaches to the Psychology of Language: Research in the Traditions of Dan Slobin. Psychology Press, London, pp. 543–555.