The Acquisition of Verb Argument Realization in Mandarin Chinese

Shu Yang and Jidong Chen

1. Introduction

The acquisition of verb argument structure has been one of the central topics in child language research. Du Bois (1987, 2003) proposed Preferred Argument Structure (henceforth PAS) as a discourse universal in adult speech to capture the statistical usage-based preference in the realization and distribution of the core grammatical arguments, S (subject of an intransitive verb) and A (subject of a transitive verb), and direct objects (O) of transitive verbs, which correlates strongly with the discourse-pragmatic factors. Recent research has shown that both adults and children are sensitive to discourse-pragmatic factors in argument realization crosslinguistically (e.g., Allen, 2009; Allen & Schröder, 2003; Clancy, 1993, 2003; Huang, 2012; Jiang & Chen, 2019; Narasimhan et al., 2005). Mandarin is a discourse-oriented language that allows extensive argument omission (e.g., Huang, 1989; Li & Thompson, 1981) and few studies have examined the development and correlation between children and adults in their realization and omission of arguments in longitudinal naturalistic speech.

This study investigates whether the argument realization in early child and caregiver speech of Mandarin Chinese (henceforth Mandarin) follows PAS (Du Bois, 1987, 2003) and the Mandarin-specific discourse-pragmatic constraints (Chui, 1992; Lin, 2009, 2011), and whether children and adults are similarly influenced by the same discourse-pragmatic and semantic factors such as information status and animacy. Two naturalistic longitudinal corpora of two monolingual Mandarin-learning children (age range 0;8 to 3;9) (Cheung et al., 2011; Deng & Yip, 2018; MacWhinney, 2000) were analyzed. A total of 8899 utterances containing the top 15 frequent verbs in the child and the caregiver speech were extracted and coded for core grammatical roles (S, A, and O), information status (given versus new), and animacy (animate versus inanimate).

2. Background

Studies of verb argument structure have focused on two dimensions: grammatical and semantical dimensions (e.g., Bresnan et al., 2015; Bresnan & Mchombo, 1987; Fillmore, 1977; Goldberg, 1995; Levin & Rappaport Hovav,

* Shu Yang, California State University, Fresno, sukkie68@mail.fresnostate.edu. Jidong Chen, California State University, Fresno, jchen@csufresno.edu.

1995; Partee, 1996). Recent research also suggested that predicate argument structure in language use could be explained pragmatically (Du Bois, 1987, 2003; Goldberg, 2001). Du Bois (1987) proposed PAS to characterize argument realization in grammatical and pragmatic dimensions based on his analyses of adult spontaneous speech in an ergative language Sakapultek Maya. The name "preferred" argument in PAS refers to arguments that are overtly realized in discourse, and "dispreferred" are those that are pronominal or omitted. In the pragmatic dimension, the information status of the arguments was classified as *given* (referents mentioned previously or notably present in the context) and *new* (referents not mentioned previously). Du Bois (1987) showed that clauses with fully realized arguments were minimal, and that no clauses contained more than one new argument. Lexical arguments (i.e., full nominal phrases) tended to express new information and appeared dominantly in the S or the O roles, whereas given information tended to be a pronoun or omitted often in the A role. The results provided evidence for "the existence and/or structural distribution of an A versus S/O structure alignment in ergative languages" (Du Bois, 1987, p. 847). Du Bois (1987) formulated four descriptive constraints in PAS to characterize argument realization in discourse, including two quantity constraints (1 and 3) and two role constraints (2 and 4): 1) the One Lexical Argument Constraint: no more than one lexical argument in each clause is preferred; 2) the Non-Lexical A Constraint: no lexical argument can appear in the A role; 3) the One New Argument Constraint: no more than one new information is preferred in each clause; and 4) the Given A Constraint: new information can only appear in the O or S role. He also suggested that the grammatical role of a referent is related to its animacy: all referents in the A role and more than 60% of referents in the S role were human, while only 10% in the O role were human referents.

PAS describes a statistical preference for argument realization in language use and has been subsequently examined and supported crosslinguistically by studies of, e.g., French and Spanish monologues (Ashby & Bentivoglio, 1993), American English conversations (Kärkkäinen, 1996), and Japanese conversations (Matsumoto, 2000). Du Bois (1987, 2003) further argued that PAS "constitutes the universally preferred locus for the introduction of new information" (Du Bois, 2003, p. 41) and could be considered a discourse universal and "argument structures are resources for speakers to exploit, for cognitive-pragmatic as well as semantic functions" (Du Bois, 2003, p. 40). The role of animacy observed in Du Bois (1987) was also attested in other languages: animate referents tend to appear in the S position in, e.g., French and Spanish (Ashby & Bentivoglio, 1993), and in American English conversations (Kärkkäinen, 1996).

A small number of studies have examined PAS in adult speech of Mandarin, a discourse-oriented language with pervasive argument omission (e.g., Huang, 1989; Li & Thompson, 1981). Long (1990) analyzed two spontaneous conversations between Mandarin-speaking adults and found evidence for all the four PAS constraints. Chui (1992) analyzed oral narratives

describing a movie by eight native Mandarin speakers and found evidence for the One Lexical Argument Constraint and the One New Argument Constraint. He also found that new referents tended to be lexical and occurred often in the O position (termed the Lexical O Constraint and the New O Constraint), and non-lexical referents occurred not only in the A but also in the S positions. Chui (1992) argued that PAS (Du Bois, 1987) could not be completely generalized to Mandarin as Mandarin speakers disfavored both the A and the S roles for new lexical arguments and the A/S alignment outnumbered the S/O alignment in Mandarin. Lin (2009, 2011) analyzed Mandarin conversations, spoken narratives, and written texts from 10 native speakers of both Mandarin and Taiwanese and found that speakers tended to produce one lexical argument per clause, and one new referent per clause, conforming to the One Argument and the One New Argument constraints (Du Bois 1987), but lexical arguments tended to be new and dominated the O role, whereas pronominals and zero arguments were often given and favored the A and the S roles. Such results support the Lexical O and the New O constraints, and the A/S alignment (Chui, 1992).

In child language research, PAS has been explored to examine children's acquisition of argument realization crosslinguistically. Supporting evidence was found in naturalistic speech of children (1;9-3;0) learning Japanese and English (Guerriero et al., 2001), Korean (1;08-2;10) (Clancy, 1993, 2003), Inuktitut (2;0-3;06) (Allen & Schröder, 2003), Hindi (2;10-4;3) (Narasimhan et al., 2005), and Tzeltal Mayan (2;0-2;5) (Brown, 1998). Narasimhan et al. (2005) argued that pragmatic prominence of arguments, as defined by information status, animacy, type of speech act, and contrastiveness (if it contrasts with other referents), crucially influenced the grammatical role and the referential form of realized arguments, and that children are sensitive to "the discourse-pragmatic factors which mediate the overt realization of arguments in their own spontaneous production" (Narasimhan et al., 2005, p. 486). Based on her analyses of Inuktitut child speech, Allen (2008) also argued that the more discourse-pragmatic factors (i.e., information status, contrastiveness, absence of referent in context, and referent differentiation in discourse) an argument is associated with, the more likely it is realized overtly, and that children are sensitive to discourse-pragmatic factors in argument realization, and an argument is more likely to be realized overtly when it is informative than when it is not.

Semantic factors such as animacy was also found to play a similar role in child speech. Clancy et al. (2003) showed that most referents in the A position were animate and the animate and inanimate referents in the S roles were almost equal, whereas the referents in the O role were inanimate primarily. Allen and Schröder (2003) also found that almost all the referents in the A role and 72% in the S role were animate, while the majority of the referents in the O roles were inanimate (69.4%).

Only a few studies (Chen, 2016; Huang, 2012, 2016; Jiang & Chen, 2019) have examined the acquisition of argument realization and PAS in Mandarin. Chen (2016) explored the emergence of verb argument structure in young

Mandarin-learning child (1;1.12 to 3;5.28) and found frequent argument omission in early child Mandarin: the majority of the transitive verbs occurred with zero or one argument, 41.3% of the intransitive verbs had no arguments, and 85% of ditransitive verbs did not sustain full arguments. Arguments of transitive verbs appeared more often in the object positions than in the subject positions. Chen (2016) did not examine the information status or the referential forms of the realized arguments from a discourse-pragmatic perspective.

Huang (2012) examined the core grammatical roles, referential forms, information status, and humanness, and clause types (transitive versus intransitive) in early naturalistic speech of two Mandarin-speaking children (2;2-3;1) in Taiwan and found evidence supporting the PAS patterns in Du Bois (1987) and Chui (1992). About 98% of the clauses contained a maximum of one argument, and only 2% had two lexical arguments. Also, 99% of the clauses had no more than one new argument. Furthermore, the A and the S roles tended to be non-lexical and given information, whereas the O role contained more lexical arguments and new referents. This pattern supports the New O and the Lexical O constraints and the A/S alignment (Chui, 1992; Lin, 2009, 2011). Moreover, the arguments in the A role were more likely to be human than in the O role. Huang (2016) further examined the PAS patterns in eight hours of conversations of two children (2;2-3;1) and their mothers and one-hour of conversations of three dyads of adult speakers from the NCCU Corpus of Spoken Chinese. She found that the One Lexical Argument and the One New Argument constraints were supported by the child, the maternal, and the adult-addressed speech, but the O role was significantly more likely to contain new and lexical referents than the A and the S roles, which supported the A/S alignment.

Jiang and Chen (2019) explored whether argument realization in elicited spoken narratives by Mandarin-English early and late bilinguals differed from those by monolingual Mandarin and English peers, focusing on the core grammatical roles (A, S, and O), referential forms, information status, and animacy. They found that all the four groups of speakers were similar in tending to avoid more than one lexical argument, lexica A, more than one new argument, or new A, which support the four PAS constraints. But the four groups of speakers differed in the degree of conformity to the four PAS constraints, which were argued to result from crosslinguistic influence, an age of second language acquisition effect, or a general effect of bilingualism.

Prior studies of Mandarin have shown that the quantity constraints of PAS (Du Bois, 1987) essentially hold in Mandarin-speaking child and adult discourse, but the new lexical referents in Mandarin tend to occur mostly in the O role, and the S and A roles are more likely to be given and non-lexical, showing the accusative S/A alignment instead of the ergative S/O alignment proposed in Du Bois (1987, 2003). No studies have examined the emergence of both realized and omitted arguments in early longitudinal child and caregiver speech considering discourse-pragmatic and semantic factors in young Mandarin-speaking children's speech.

3. The current study

Our specific research questions are:

(1) Do the argument realization and omission in early child speech of Mandarin follow PAS (Du Bois, 1987, 2003) and the Mandarin-specific discourse-pragmatic constraints (e.g., Chui, 1992)?
(2) Are children and adults are similarly influenced by the same discourse-pragmatic factors such as information status and animacy in argument realization and omission?

3.1. Methodology

We examined two longitudinal naturalistic corpora of two children (age range 0;8 to 3;9) (see Table 1), Tong from the Tong Mandarin corpus (Deng & Yip, 2018; MacWhinney, 2000), and Pan from the Taiwan Corpus of Mandarin Chinese (TCMC) (Cheung et al., 2011; MacWhinney, 2000). The Tong corpus includes 22 transcripts of monthly one-hour recordings of the naturalistic interactions between a Mandarin-learning child Tong and his caregivers (mother, father, and grandparents) from 1;7 to 3;4. The Pan corpus contains 19 transcripts of bimonthly recordings of the naturalistic interactions between the child Pan and the investigator and his caregivers (mother, father, and grandmother) from 1;7 to 3;9. Both corpora contain morphosyntactic coding tiers.

Table 1. Background information of the target children in the corpora

	Age range	No. Files	No. Child Utterances	No. of adult utterances	Types	Tokens
Tong	1;07-3;04	22	9,111	21,284	1,848	3,0813
Pan	1;07-3.09	19	5,322	6,667	1,023	8,578

All the verbs in the corpora were extracted using the Computerized Language Analysis (CLAN) program (MacWhinney, 2000). A total of 441 types of verbs were produced by Tong and 390 by Pan. In the current study the top 15 frequent verbs from each child's speech (see Appendix) were included in the analyses. All the main arguments of those 15 frequent verbs in the children's and the caregivers' speech were coded for the following features based on the coding categories in Allen (2008) and Chen (2016): 1) verb type (transitive, intransitive, and ditransitive verbs), 2) argument types: S, A, Direct Object (DO), and Indirect Object (IO), 3) number of lexical arguments, 4) argument omission type (subject, object, and no omission), and 5) information status of arguments (new versus given), and animacy of arguments (animate versus inanimate). Omitted arguments were identified through discourse context and were coded for the same categories of features above. A total of 8899 sentences were coded, including 6449 in the Tong corpus and 2450 utterances in the Pan corpus. We

excluded incomplete and unintelligible sentences (50), imperative sentences (1,103), existential constructions with the verb *yǒu* 'to have' (1,523), and sentences with the verb *yào* 'will' as model verb (752).

3.2. Results

Table 2 presents the proportions of the clauses with different number of arguments in Pan, Tong, and their caregiver speech. It shows that both children produced utterances with a maximum of one argument (Pan 72% and Tong 70%) more often than those with two (27% and 29% respectively) or three arguments (1% for both). This distribution resembles that in the speech of their caregivers: zero or one argument (61% and 74% in Pan and Tong input), two (38% and 25% respectively) or three arguments (1% for both).

Table 2. Distribution of clauses with different number of the arguments

Argument no.	Pan	Pan's input	Tong	Tong's input
0	29%	20%	20%	23%
1	43%	41%	50%	51%
2	27%	38%	29%	25%
3	1%	1%	1%	1%

To investigate if the PAS patterns in Du Bois (1987) and Chui (1992) are supported by the Pan and the Tong corpora, we examined each constraint below.

The One Lexical Argument Constraint. The One Lexical Argument Constraint suggests that no more than one lexical argument is preferred in each clause (e.g., Du Bois, 1987, 2003). Table 3 shows the percentages of the utterances with different number of lexical arguments in Pan, Tong, and their caregiver speech.

Table 3. Distribution of clauses with different number of lexical arguments

Lexical argument	Pan	Pan's input	Tong	Tong's input
0	54%	59%	44%	46%
1	39%	36%	48%	43%
2	7%	4%	7%	10%
3	0%	1%	0%	1%

As shown in Table 3, more than 90% of both Pan's and his caregivers' speech contain zero or one lexical argument: zero lexical arguments (54% in Pan's and 59% in his caregivers' speech) and one lexical argument (39% and 36% respectively); utterances with two lexical arguments are rare (7% and 4% respectively) and three lexical arguments (0% and 1% respectively). The same pattern is seen in Tong's data: utterances with zero or only one lexical argument predominate (92% for Tong and 89% for Tong's caregivers) and utterances with more than one lexical argument is minimal (7% and 11% in Pan's and his

caregivers' speech respectively). The results thus conform to the One Lexical Argument Constraint.

The Non-Lexical A Constraint. The Non-Lexical A Constraint indicates that lexical argument is avoided at the A role. Table 4 shows the distribution of lexical and non-lexical arguments across different grammatical roles.

Table 4. Distribution of lexical versus non-lexical arguments by grammatical roles (L=lexical, NL=non-lexical)

	Pan		Pan's input		Tong		Tong's input	
	L	NL	L	NL	L	NL	L	NL
S	35%	65%	25%	75%	16%	84%	18%	82%
A	13%	87%	14%	86%	9%	91%	22%	78%
DO	47%	53%	41%	59%	67%	33%	56%	44%
IO	20%	80%	31%	69%	72%	28%	78%	22%

As shown above, most of the A roles contain non-lexical arguments for both Pan (87%) and Tong (91%) and their respective caregivers (86% and 78%). The results thus support the Non-Lexical A Constraint.

The Lexical O Constraint. The Lexical O Constraint (e.g., Chui, 1992) characterizes the distribution in Mandarin that lexical arguments usually appear as the objects of transitive verbs. This constraint is supported only by Tong's data: more lexical O's occur in the speech of Tong (67% DO's and 72% IO's) and his caregivers (56% DO's and 78% IO's). Pan, on the other hand, shows almost an equal distribution of lexical (47%) and non-lexical (53%) O's, and lexical IO's comprises only 20% IO's. Pan's divergence from the Lexical O constraint may have resulted from the similar dominance of non-lexical O's (59% DO's and 69% IO's) in his caregivers' speech. Note that the lexical arguments in the O role (DO and IO) in the speech of Pan and his caregivers are still relatively higher than the lexical arguments in the A and S role, suggesting that O's are overall more likely to be lexical than the A and the S arguments.

The One New Argument Constraint. The One New Argument Constraint suggests that no more than one new argument is preferred in each clause. Table 5 summarizes the distribution of the utterances with different number of new arguments in the data. It shows that utterances with no or one new argument predominate in the speech of both children and their caregivers (about 99%). This pattern is consistent with the One New Argument Constraint.

Table 5. Distribution of sentences with new arguments per clause

new argument per clause	Pan		Pan's input		Tong		Tong's input	
	Frequency (%)		Frequency (%)		Frequency (%)		Frequency (%)	
0	242	(71%)	855	(58%)	665	(76%)	1882	(79%)
1	96	(28.2%)	609	(41.4%)	208	(23%)	475	(20%)
2	3	(0.8%)	9	(0.6%)	11	(1%)	28	(1%)
Total	341	(100%)	1473	(100%)	874	(100%)	2385	(100%)

The Given A Constraint and the New O Constraint. The Given A Constraint (Du Bois, 1987) suggests that old information often occurs in the A role and new information can only appear in the O or S role. The New O Constraint (e.g., Chui, 1992) argues that in Mandarin the new arguments occur more frequently in the O role, whereas the S and the A roles rarely contain new referents. As Table 6 shows, both children (Pan 97% and Tong 99%) and their caregivers (99% for both) show a dominance of given A's, and new referents also occur more frequently as DO's than A's in the speech of both Pan and Tong (30% each) and their caregivers (45% and 26% respectively), which are consistent with the Given A Constraint. But new arguments rarely occur in the S role for both children (3% Pan and 0% Tong) and their caregivers (2% and 1% respectively), which is consistent with the New O Constraint. However, almost all IO's are given in the speech of Pan (100%) and his caregivers (94%), whereas the majority (59%) of IO's in Tong's speech contain new referents, which may result from a higher rate of new IO's in his caregivers' speech (45%).

Table 6. Distribution of arguments by information status

	Pan		Pan's input		Tong		Tong's input	
	New	Given	New	Given	New	Given	New	Given
S	3%	97%	2%	98%	0%	100%	1%	99%
A	3%	97%	1%	99%	1%	99%	1%	99%
DO	30%	70%	45%	55%	30%	70%	26%	74%
IO	0%	100%	6%	94%	59%	41%	45%	55%

PAS and the omitted arguments. The analyses above examined the realized arguments, and this section reports the results of the omitted arguments in terms of their discourse-pragmatic features. Overall, the majority of the sentences produced by both children omitted at least one argument (68% for Pan and 56.2% for Tong), similar to their caregivers' production (59% and 56.4% for Pan's and Tong's caregivers respectively). Figure 1 presents the distribution of the realized and the omitted arguments by grammatical roles in the speech of the two children and their caregivers. As shown in Figure 1, subjects (A and S) are omitted more frequently than objects (DO and IO) in Pan's data: 52% and 56% omitted S and A arguments versus 42% and 0% omitted DO and IO arguments, and 37% and 45% omitted S and A arguments versus 39% and 0% omitted DO and IO arguments in his caregivers' speech. A similar pattern is found in Tong's data: 40% and 56% omitted S and A arguments versus 22% omitted DO and 4% omitted IO arguments, and 44% and 48% omitted S and A arguments versus 35% and 6% omitted DO and IO arguments in his caregivers' speech.
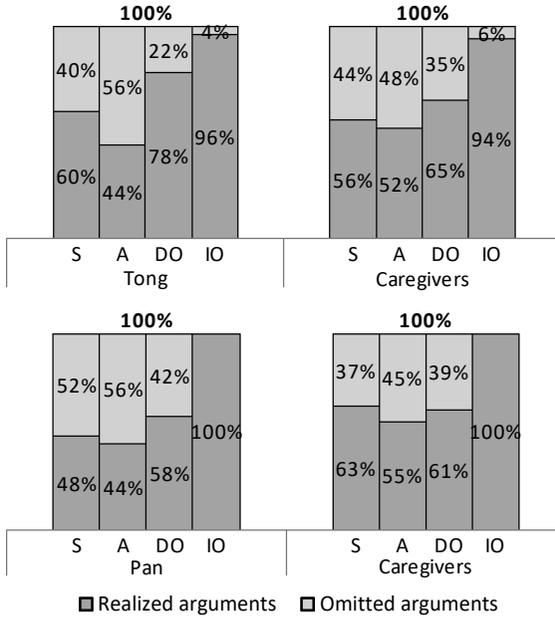
Figure 1. Distribution of overt and omitted arguments by grammatical role

The information status of the arguments is further examined to reveal how discourse-pragmatic factors affect argument omission. Table 7 shows the proportions of the omitted arguments by information status in the data. There is a strong preference to omit given arguments across the children and their caregivers regardless of the grammatical roles, ranging from 82% to 100%. Given S, A, and IO arguments are omitted more often (94% and above) than given DO arguments in the speech of both the children (82% for Pan and 90% for Tong) and their caregivers (82% and 90% respectively).

Table 7. Distribution of omitted arguments by information status

	Pan		Pan's input		Tong		Tong's input	
	New	Given	New	Given	New	Given	New	Given
S	6%	94%	0%	100%	0%	100%	0%	100%
A	2%	98%	0%	100%	0.4%	99.6%	1%	99%
DO	18%	82%	18%	82%	10%	90%	10%	90%
IO	6%	94%	0%	100%	0%	100%	8%	92%

The distribution of omitted arguments by animacy is shown in Table 8. Almost all the omitted subjects (S and A) are animate referents (96%-100%) and most of the elided DO's are inanimate referents (91%-97%) in the speech of both the children and their caregivers. Only a small amount of the elided DO's are animate referents in the speech of both the children and their caregivers

(2.4%-9%). Pan and his caregivers did not omit any IO arguments and Tong and his caregivers omitted mostly inanimate IO arguments (97%-100%). To sum up, subjects are more often omitted than objects and given arguments are usually omitted by both the children and the caregivers. Almost all the omitted subjects (S and A) are given and animate referents, and most of the omitted DO's are given and inanimate referents in the child speech and the input.

Table 8. Distribution of omitted arguments by animacy (A: Animate; IA: Inanimate)

	Pan		Pan's input		Tong		Tong's input	
	A	IA	A	IA	A	IA	A	IA
S	100%	0%	100%	0%	100%	0%	100%	0%
A	96%	4%	100%	0%	100%	0%	99.6%	0.4%
DO	9%	91%	3%	97%	2.4%	97.6%	3%	97%
IO	0	0	0	0	0%	100%	3%	97%

4. Discussion and conclusion

We investigated whether emergent argument realization and omission in early child (0;8 to 3;09) and caregiver speech of Mandarin is influenced by discourse-pragmatic and semantic factors such as information status and animacy of the arguments, and whether children and adults are similarly influenced by these factors. Our results show that child and adult speakers of Mandarin follow the PAS constraints (Du Bois, 1987), but new and lexical arguments occur more often in the O role and both the A and the S roles tend to be given and non-lexical, supporting the A/S alignment (e.g., Chui, 1992). The patterns of argument realization and omission in the speech of both the children and their caregivers are similarly conditioned by the information status and the animacy of the arguments.

PAS in Mandarin child and caregiver speech. Consistent with the One Lexical Argument and the One New Argument constraints (Du Bois, 1987), most of the utterances from the children and the caregivers contain a maximum of one lexical and one new argument. This result confirms prior research of production of argument structure in Mandarin-speaking adults (Chui, 1992; Lin, 2009; Long, 1990), monolingual Mandarin children (Chen, 2016; Huang, 2012, 2016), and bilingual Mandarin speakers (Jiang & Chen, 2019). Partially consistent with the Non-Lexical A Constraint and the Given A Constraint, non-lexical arguments are preferred in not only the A but also the S roles, and old information mostly appears in not only the A but also S roles. These results thus support the A/S alignment in Mandarin (Chui, 1992; Huang, 2012; Lin, 2009, 2011) and differs from the S/O alignment in ergative languages such as Sakapultek (Du Bois, 1987, 2003).

Regarding the features of the objects, the Lexical O and the New O constraints (Chui, 1992) are supported by the distribution of the DO's by both children and their caregivers. Consistent with the New O Constraint (Chui,

1992), new referents occur more frequently as DO's than A's in the speech of both children and their caregivers, even though one of the children, Pan, shows almost an equal distribution of lexical and non-lexical O's. The distribution of the IOs supports the Lexical O and the New O constraints only by one of the children (Tong) and his caregivers: the majority of the IOs in Tong and his caregivers' speech are new and lexical referents, but almost all the IOs in the Pan corpus are given with only 20% being lexical. Such results reveal individual differences in argument realization and acquisition, and reflect the influence of different distributions in the input.

Factors influencing the argument realization and omission. Information status and animacy of the referents have been found to influence argument realization and omission in child language, for example, given information is often omitted and new information tends to be maintained (Allen & Schröder, 2003; Clancy, 1993, 1997, 2003; Huang, 2012, 2016; Jiang & Chen, 2019). Our study shows that all the omitted arguments in the children and their caregivers' speech across the grammatical roles are given referents, and the realized arguments tend to be new. Animacy of the referents plays a role in the realization or omission of verb arguments (e.g., Everett, 2009; Huang, 2012). Mandarin-speaking children are found more likely to have human referents in the A and the S roles and often omit them or use the pronominal form *nǐ* "I" and *wǒ* "you" (Huang, 2012). Our result shows that nearly all the overt and omitted subjects (A and S) contain animate referents, and all the omitted subjects are first- or second-person referents. Most of the overt and omitted objects are inanimate and third-person referents. Our study thus provides further evidence that the argument realization in early child language can be discourse-motivated and children are sensitive to the information status and semantics of the referents in deciding on the referential choice in their communication (Allen, 2008; Huang, 2012, 2016).

Relationship between the child output and input. Prior studies have shown that the acquisition of the verb argument realization is closely related to the distributional properties of adult uses of verb argument patterns in naturalistic conversations with children (Chen, 2016; Clancy, 1997; Huang, 2012, 2016). Our results reveal that both the argument realization and the argument omission in the two children's speech closely mirror the usage of their caregivers', and that individual variation in the child's production can be accounted for by the same distributional patterns in their respective caregivers' production. Narasimhan et al. (2005) argue that input provides a relatively direct route for children to realize the verb transitivity. Such close correlation and distributional homogeneity between the input and the output have been widely documented in child language acquisition research of a diverse range of constructions (Behrens, 2006; Lieven et al., 2003; Slobin, 1991, 1997; Tomasello, 2001; Valian, 1991; Wijnen et al., 2001).

To conclude, our study provides further evidence that Mandarin-acquiring children are sensitive to the discourse-pragmatic factors in argument realization from an early age, and the PAS constraints (Du Bois 1987, 2003) are generally

supported by child Mandarin and their input, but the results are consistent with the A/S alignment instead of the S/O alignment (cf. Chui, 1992; Ling, 2009, 2011; Huang, 2012). The speech of both the child and their caregivers is modulated by the same discourse-pragmatic factors, i.e., information structure and animacy of the referents. Individual differences between the children reflect the different usage patterns in their respective input. Our study supports the discourse-pragmatic and usage-based probabilistic approaches to language and language acquisition and contributes to the theoretical and empirical studies of the acquisition argument structure crosslinguistically.

Appendix: Top 15 most frequent verbs by child (shared verbs in bold)

	Tong		Token	Pan		Token
1	**yào**	**"to want"**	**813**	**yào**	**"to want"**	**282**
2	**yǒu**	**"to have"**	**644**	**yǒu**	**"to have"**	**192**
3	lái	"to come"	478	chī	"to eat"	66
4	**kàn**	**"to look"**	299	zuò	"to do"	42
5	**qù**	**"to go"**	277	gěi	"to give"	39
6	kāi	"to open"	205	**qù**	**"to go"**	**39**
7	**dǎ**	**"to hit"**	**176**	xiào	"to laugh"	32
8	zǒu	"to walk"	163	**kàn**	**"to look"**	**30**
9	shàng	"to get on"	148	**zhīdào**	**"to know"**	**24**
10	**wán**	**"to play"**	**139**	**wán**	**"to play"**	**23**
11	**zhīdào**	**"to know"**	**133**	**fàng**	**"to put"**	**22**
12	zuò	"to sit"	130	**ná**	**"to take"**	**21**
13	xià	"to get off"	128	dài	"to bring"	21
14	**fàng**	**"to put"**	**112**	**dǎ**	**"to hit"**	**21**
15	**ná**	**"to take"**	109	diū	"to throw"	21

References

Allen, Shanley E. M. (2008). Interacting pragmatic influences on children's argument realization. In M. Bowerman & P. Brown (Eds.), *Crosslinguistic perspectives on argument structure: Implications for learnability* (pp. 191-210). Taylor & Francis.

Allen, Shanley E. M. (2009). Argument structure. In E. L. Bavin (Ed.), *Cambridge handbook of child language* (pp. 217-236). Cambridge University Press.

Allen, Shanley E. M., & Schröder, Heike. (2003). Preferred argument structure in early Inuktitut spontaneous speech data. In J. W. Du Bois, L. E. Kumpf, & W. J. Ashby (Eds.), *Preferred argument structure: Grammar as architecture for function* (Vol. 14, pp. 301-338). John Benjamins Publishing.

Ashby, William J, & Bentivoglio, Paola. (1993). Preferred argument structure in spoken French and Spanish. *Language variation and change*, *5*(1), 61-76. https://doi.org/10.1017/S095439450000140X

Behrens, Heike. (2006). The input–output relationship in first language acquisition. *Language and Cognitive Processes*, *21*(1-3), 2-24.

Bresnan, Joan, Asudeh, Ash, Toivonen, Ida, & Wechsler, Stephen. (2015). *Lexical-functional syntax*. John Wiley & Sons.

Bresnan, Joan, & Mchombo, Sam A. (1987). Topic, pronoun, and agreement in Chicheŵa. *Language*, *63*, 741-782. https://doi.org/10.2307/415717

Brown, Penelope. (1998). Conversational structure and language acquisition: The role of repetition in Tzeltal. *Journal of Linguistic Anthropology*, *8*(2), 197-221. https://doi.org/10.1525/jlin.1998.8.2.197

Chen, Jidong. (2016). The emergence of verb argument structure in Mandarin Chinese. . In H. Tao (Ed.), *Integrating Chinese linguistic research and language learning and teaching* (pp. 1-11). John Benjamins.

Cheung, Hintat, Chang, Chien-Ju, Ko, Hwa-Wei, & Tsay, Jane. (2011). *Taiwan corpus of child Mandarin (TCCM)*. The CHILDES database. https://childes.talkbank.org/access/Chinese/Mandarin/TCCM.html

Chui, Ka-Wai. (1992). Preferred argument structure for discourse understanding. *The Proceedings of COLING 1992*, 1142-1146.

Clancy, Patricia M. (1993). Preferred argument structure in Korean acquisition. In E. E. Clark (Ed.), *Proceedings of the 25th annual child language research forum* (Vol. 25, pp. 307-314). CSLI Publications.

Clancy, Patricia M. (1997). Discourse motivations for referential choice in Korean acquisition. In H. Sohn & J. Haig (Eds.), *Japnanese/Korean Linguistics* (Vol. 6, pp. 639-659). Center for the Study of Language and Information Publications.

Clancy, Patricia M. (2003). The lexicon in interaction: Developmental origins of preferred argument structure in Korean. In J. W. Du Bois, L. E. Kumpf, & W. J. Ashby (Eds.), *Preferred argument structure: Grammar as architecture for function* (pp. 81-108). John Benjamins.

Deng, Xiangjun, & Yip, Virginia. (2018). A multimedia corpus of child Mandarin: The Tong corpus. *Journal of Chinese Linguistics*, *46*(1), 69-92.

Du Bois, John W. (1987). The discourse bais of ergativity. *Language 63*(4), 805-855. https://doi.org/10.2307/415719

Du Bois, John W. (2003). Argument structure: Grammar in use. In J. W. Du Bois, L. E. Kumpf, & W. J. Ashby (Eds.), *Preferred argument structure: Grammar as architecture for function* (Vol. 14, pp. 1-60). John Benjamins Publishing.

Everett, Caleb. (2009). A reconsideration of the motivations for preferred argument structure. *Studies in Language*, *33*(1), 1-24. https://doi.org/10.1075/sl.33.1.02eve

Fillmore, Charles J. (1977). The case for case reopened. In P. Cole & J. M. Sadock (Eds.), *Grammatical relations* (pp. 59-81). Brill.

Goldberg, Adele E. (1995). *Constructions: A construction grammar approach to argument structure*. University of Chicago Press.

Goldberg, Adele E. (2001). Patient arguments of causative verbs can be omitted: The role of information structure in argument distribution. *Language Sciences*, *23*(4-5), 503-524. https://doi.org/10.1016/S0388-0001(00)00034-6

Guerriero, Sonia, Cooper, Amy, Oshima-Takane, Yuriko, & Kuriyama, Yoko. (2001). A discourse-pragmatic explanation for argument realization and omission in English and Japanese children's speech. In A. H.-J. Do, L. Domínguez, & A. Johansen (Eds.), *Proceedings of the 25th annual Boston university conference on language development* (Vol. 1, pp. 319-330). Cascadilla Press.

Huang, C. T. James. (1989). Pro-drop in Chinese: A generalized control theory. In O. A. Jaeggli & K. J. Safir (Eds.), *The null subject parameter* (pp. 185-214). Springer.

Huang, Chiung-chih. (2012). Preferred argument structure in Mandarin child language. *Taiwan Journal of Linguistics*, *10*(2), 119-168.

Huang, Chiung-Chih. (2016). Information management in Mandarin child speech, maternal speech, and adult speech. *Lingua*, *184*, 53-68. https://doi.org/10.1016/j.lingua.2016.06.001

Jiang, Xiangyu, & Chen, Liang. (2019). Preferred argument structure in the narratives of Chinese-English bilinguals and their monolingual peers. *International journal of bilingualism*, *23*(5), 873-887. https://doi.org/10.1177/1367006917728397

Kärkkäinen, Elise. (1996). Preferred argument structure and subject role in American English conversational discourse. *Journal of Pragmatics*, *25*(5), 675-701.

Levin, Beth, & Rappaport Hovav, Malka. (1995). *Unaccusativity: At the syntax-lexical semantics interface*. MIT press.

Li, Charles, & Thompson, Sandra. (1981). *Mandarin Chinese: A functional reference grammar*. University of California Press.

Lieven, Elena, Behrens, Heike, Speares, Jennifer, & Tomasello, Michael. (2003). Early syntactic creativity: A usage-based approach. *Journal of Child Language*, *30*(2), 333. https://doi.org/10.1017/S0305000903005592

Lin, Wan-Hua. (2009). Preferred argument structure in Chinese: A comparison among conversations, narratives, and written texts. In Y. Xiao (Ed.), *The proceedings of the 21st North American conference on Chinese linguistics* (Vol. 2, pp. 341-357). Smithfield, Rhode Island: Bryant University.

Lin, Wan-Hua. (2011). Preferred argument structure in Chinese conversations, narratives, and written texts. In Y. Xiao, L. Tao, & H. L. Soh (Eds.), *Current issues in Chinese linguistics* (pp. 332-363). Cambridge Scholars Publishing.

Long, Zhihua. (1990). The Preferred Argument Structure in Mandarin Chinese and its cognitive implications. *Colorado Research in Linguistics*, *11*, 42-49.

MacWhinney, Brian. (2000). *The CHILDES project: Tools for analyzing talk* (3rd ed.). Lawrence Erlbaum.

Matsumoto, Kazuko. (2000). Intonation units, clauses and preferred argument structure in conversational Japanese. *Language Sciences*, *22*(1), 63-86. https://doi.org/10.1016/S0388-0001(99)00004-2

Narasimhan, Bhuvana, Budwig, Nancy, & Murty, Lalita. (2005). Argument realization in Hindi caregiver–child discourse. *Journal of Pragmatics*, *37*(4), 461-495. https://doi.org/10.1016/j.pragma.2004.01.005

Partee, Barbara H. (1996). The development of formal semantics in linguistic theory. In S. Lappin (Ed.), *The handbook of contemporary semantic theory* (pp. 11-38). Blackwell.

Slobin, Dan Isaac. (1991). Learning to think for speaking: Native language, cognition, and rhetorical style. *Pragmatics*, *1*(1), 7-25. https://doi.org/10.1075/prag.1.1.01slo

Slobin, Dan Isaac. (1997). The universal, the typological, and the particular in acquisition. In D. I. Slobin (Ed.), *The crosslinguistic study of language acquisition.* (Vol. 5: Expanding the contexts, pp. 1-39). Lawrence Erlbaum.

Tomasello, Michael. (2001). First steps toward a usage-based theory of language acquisition. *Cognitive Linguistics*, *11*(1-2), 61-82. https://doi.org/10.1515/cogl.2001.012

Valian, Virginia. (1991). Syntactic subjects in the early speech of American and Italian children. *Cognition*, *40*(1-2), 21-81. https://doi.org/10.1016/0010-0277(91)90046-7

Wijnen, Frank, Kempen, Masja, & Gillis, Steven. (2001). Root infinitives in Dutch early child language: an effect of input? *Journal of Child Language*, *28*(3), 629. https://doi.org/10.1017/S0305000901004809

Proceedings of the 46th annual Boston University Conference on Language Development

edited by Ying Gong
and Felix Kpogo

Cascadilla Press   Somerville, MA   2022

Ordering information

To order a copy of the proceedings or to place a standing order, contact: