The Emergence of Verb Argument Structure in Mandarin Chinese

Jidong Chen

California State University, Fresno

This study examines the emergent linguistic properties of the early production of verb argument structure in Mandarin Chinese. Longitudinal naturalist speech data of one Mandarin-learning child were coded and analyzed for their argument structure from age 1;3 to 3;4, the crucial age for early syntactic development. The results reveal early emergence and dominance of transitive verbs and the theme-only argument. Transitive verbs tend to have reduced number of arguments, whereas intransitive verbs tend to sustain one argument. The acquisition of thematic roles proceeds in the order of theme, agent, location, and recipient. The child's production of verb argument structure reflects the typological features of Mandarin. It also suggests some language-general route in the acquisition of argument structure.

0. Introduction

To use a verb grammatically, a child needs to figure out how linguistically event participants are realized syntactically, i.e. the verb argument structure. Syntactically an argument is defined as 'a noun phrase bearing a specific grammatical or semantic relation to a verb and whose overt or implied presence is required for well-formedness in structures containing that verb' (Trask 1993: 20). An argument structure typically indicates the number of arguments (Noun Phrases) a verb takes (e.g., the core participants in the eventuality a verb denotes), their syntactic expression, and their semantic relation to this lexical item. Arguments can be identified in two ways: in terms of syntactic roles with respect to the verb such as Subject and Object, and in terms of semantic roles in relation to the verb such as Agent (entity that initiates an action) and Patient (entity that undergoes an action). The mapping between the semantic and the syntactic roles and the conceptual representation of a real-world event is not a straightforward one. For example, it is ungrammatical to say **John laughed Mary*¹ to describe an event in which John made Mary laugh, even though it is what a child may have observed or recorded in his or her mental representation. In other words, to use a verb grammatically, a child has to acquire not only the semantics of the verb correctly but also the syntactic structure that his or her language allows this verb to occur in.

Mandarin presents an interesting case for the study of verb argument structure. It has a number of important typological features that distinguish it from many European languages such as English or German. It is a discourse-oriented language (e.g. Li &

---

¹ The asterisk * indicates ungrammaticality.

Thompson 1981) that allows extensive argument omission when retrievable from the context. It is also morphologically impoverished – there is no morphological marking to indicate the syntactic or semantic roles of the arguments when arguments are omitted.

Much of the literature on argument structure acquisition assumes that arguments are always provided in caregiver speech to present the full input necessary to learners. However, this is probably not the case in pro-drop language such as Mandarin. This scenario thus presents a worse learning situation than the classical ‘poverty-of-stimulus’ problem noted in Chomsky’s (1999) theory of language acquisition. The omission of arguments thus poses two challenges for children: How do children receive sufficient data from the input to learn argument structure, and how do they display their knowledge of argument structure?

1. The current study

The current study focuses on two specific research questions: (1) What do the L1 data reveal about the emergence of verb argument structure in Mandarin? And (2) To what extent does language-specific morph-syntactic properties affect the path of the acquisition of argument structure?

1.1 Data

To answer these research questions about the developmental trajectory of verb production, I analyzed naturalistic longitudinal speech data of a Mandarin-learning child MDY (ages from 1;01.12 to 3;5.28) from the Fang corpus (Chen 2008; Min 1994). MDY was visited and audio-taped at home in Beijing once a week or once every other week from 1984 to 1986. The data were mostly collected in natural situations at the target child’s home, and the sessions all involved multiple adult caregivers interacting with the child, as is typical for Chinese children of their age. The have been transcribed in *Pinyin*, the official Mandarin Romanization system used in China, and followed the transcription format of the CHAT (Codes for the Human Analysis of Transcripts) used by the CHILDES (MacWhinney 2000). A total of 16 sessions in MDY’s corpus from ages 1;3 to 3;4 were included for analysis based on the criterion that the child had to be at least 1;0 at the time of the first file selected, by which age we could expect utterances of one word or longer.

Among the total of 2645 utterances produced by MDY, 1358 (51%) contained a verbal predicate. Adjectival predicates were included in the coding and analysis as Mandarin adjectives are known to function fully as a predicate (Li & Thompson 1981). To establish the child’s overall morphological and syntactic development, I calculated the MLU (Mean Length of Utterance) using The CLAN (Computerized Language Analysis) program (MacWhinney 2000). Mean Length of Utterance (MLU) has been widely used in language acquisition studies around the world (e.g. Brown 1973). This measure has been argued to be a valid and useful tool for Mandarin as well (Cheung 1998), and it has been used in major studies of the acquisition of Mandarin (Erbaugh 1978; Erbaugh 1982;

Erbaugh 1992; Min 1994; Tardif 1993). The sessions were also grouped to represent different developmental stages of development measured by MLU. Table 1 presents the details of the MLUs, the numbers and proportions of verbal clauses of each session, and the mean MLUs of each developmental stage. At stage 1 the child's mean MLU is 1.75 between ages 1;3 to 1;5. At stage 2 the mean MLU increases to 2.99 at the age of 1;7 to 2;7. At stage 3 the mean MLU reaches 4.00 at the age of 3;0-3;5. Note that the division of data is for the convenience of visualizing the development and it does not assume that the development is discontinuous.

Table 1. Overview of the data: Numbers and proportions of verbal clauses and the three developmental stages measured by MLUs

Age	No. of utterances	No. of morphemes	MLU	No. of verbal clauses coded	Percentages of verbal clauses	
1;3.15	180	240	1.333	34	19%	Stage 1: MLU = 1.75 (1;3-1;5)
1;4.6	128	206	1.609	34	27%	
1;5.30	164	345	2.104	42	26%	
1;7.9	209	415	1.986	92	44%	Stage 2: MLU = 2.99 (1;7-2;7)
1;10.13	82	228	2.78	41	50%	
1;11.12	91	118	2.066	37	41%	
2;0.0	243	722	2.971	118	49%	
2;1.03	119	366	3.076	72	61%	
2;2.27	91	246	2.703	35	38%	
2;3.6	143	506	3.538	84	59%	Stage 3: MLU = 4.00 (3;0-3;5)
2;7.3	58	169	2.914	31	53%	
3;0.5	219	732	3.342	87	40%	
3;1.10	122	468	3.836	74	61%	
3;2.3	135	551	4.081	96	71%	
3;3.4	330	1368	4.145	232	70%	
3;4.15	331	1492	4.508	249	75%	
Total	2645	8172		1358	51%	

1.2 Coding

All the verbal utterances in the child's speech were coded for the transitivity of the verb, the number of arguments (e.g. 1 argument), the thematic roles of the arguments (e.g. Agent), and the syntactic role in the clause (e.g. subject). Sample sentences with their coding are illustrated in Table 2.

Table 2. Sample sentences and their coding

One-argument sentences		
Examples	Thematic roles	Syntactic roles
我吃了。 <i>wo chi le.</i> 'I eat PFV' (I ate)	Agent	Subj
小熊醒了。 <i>xiaoxiong xing le.</i> 'little.bear woke.up PFV' (Little bear woke up)	Theme	Subj
Two argument sentences		
她吃了苹果。 <i>ta chi le pingguo.</i> 'he eat PFV apple' (he ate an apple)	Agent-Theme	Subj-obj
球球滚到那儿了。 <i>qiú qiú gǔn dào nàr le.</i> 'ball roll to there PFV' (the ball rolled there)	Theme-Location	Subj-obj
Three-argument sentences		
她给我苹果。 <i>ta gei wo pingguo.</i> 'he give me apple' (he gave me apples)	Agent-Recipient-Theme	Subj-obj-obj

1.3 Results

Overall the most frequent type of verbs is transitive verbs (67.9%), followed by intransitive verbs (15.8%), adjectival predicates (7.9%), ditransitive verbs (4.5%), and copular verbs (3.9%). The majority of the sentences contain only 1 argument (51.4%). Sentences with null arguments (26.7%) and 2 arguments (21%) are also common, and sentences with 3 arguments are quite rare (less than 1%). The contrast between the dominance of transitive verbs and 1-argument sentences suggests that argument omission is overall frequent.

To see how arguments are realized across different verb types, I calculated the proportions of argument structure types within each verb type, i.e. transitive verbs, intransitive verbs, and ditransitive verbs. Figure 1 shows the distribution of argument structures across the three types of verbs. Within transitive verbs, 1-argument sentences are the most frequent (49.7%), followed by null-argument sentences (23.3%). Only 26.6% of transitive verbs have full realization of both arguments. Within intransitive verbs, the majority of the sentences sustain 1 full argument (58.7%) and the rest have no arguments (41.3%). Within ditransitive verbs, 1-argument sentences are the most common (58.3%), followed by 2-argument structures (26.7%). Only 15% of the ditransitive verbs have full 3 arguments. This pattern suggests that transitive and ditransitive verbs tend to have reduced number of arguments whereas intransitive verbs are more likely to sustain one argument.

Figure 1. Overall distribution of argument structure types by verb type

It is well recognized in the child language acquisition literature that argument omission is common in early child speech in pro-drop languages such as Mandarin and non-pro-drop languages such as English. An asymmetry has also been widely observed in syntactic positions of the omitted arguments, i.e. subjects being omitted more often than objects in a number of typologically diverse languages (Allen 2000; Hyams & Wexler 1993). To find out if such an asymmetry also occurs in Mandarin, I calculated the proportions of the arguments at different syntactic positions when they are produced. It turned out that a single argument occurs equally likely in either the subject or the object positions (48%) across all the verb types. But a further examination shows an asymmetry in the occurrence of arguments syntactically for transitive verbs: arguments are more likely to occur in the object position (40%) than in the subject positions (23.1%). This pattern is in line with the widely recognized asymmetry that subjects are omitted more often than objects.

All the verb arguments were also examined for the semantic roles. A variety of thematic roles were produced, including theme, agent, recipient, and location. The most frequent type of thematic relations is theme-only (45.3%), followed by agent-only (21.5%) and agent-theme (19.7%). This pattern indicates that theme-oriented argument structures are the most productive.

Let us move on to the developmental trajectory in the child's production of verb argument structure. As explained earlier in the coding section, the data sessions were grouped to represent different stages of development measured by MLUs. Each stage was analyzed for the proportions of the token frequencies of different types of argument structures and their semantic roles. Figure 2 shows the distribution of argument structure types across the three stages. Sentences with only 1 argument are the most frequent at stage 1 (61.3%) and remain the most frequent across all three stages, 48.6% at stage 2 and 51.8% at stage 3, despite some decrease. Sentences with no arguments comprise a

big portion, 37% at stage 1, and decrease to 21.5% at stage 3. Sentences with 2 arguments comprise about 21% of all the verbal utterances at stages 1 and 2, and increase slightly at stage 3 to about 25%. Sentences with 3 arguments are overall very rare across the stages. This pattern suggests that argument omission is frequent from the earliest stage onward, and sustaining 1 argument is preferred across all the stages.

Figure 2. Development of argument structure types by argument number

In terms of the thematic roles, theme, agent, and location are among the earliest, and recipient appears at stage 2 (5-6%). Theme-only sentences predominate stage 1 (86%) and remain the most frequent across all the stages despite a sharp decrease to 49.3% and 41% at stages 2 and 3 respectively, as illustrated in Figure 3. Agent-only sentences represent a small portion at stage 1 (7.8%) and increase gradually from 7.8% at stage 2 to 26.2% at stage 3. Agent-theme sentences are rare at stage 1 and grow to 16.5% at stage 2 and 21.8% at stage 3. Thus the theme-oriented structure emerges earlier and is more productive than agent and other thematic roles.

Figure 3. Development of argument structure by thematic roles

2. Discussion and conclusion

This study presents a case study of a Mandarin-learning child's longitudinal production of verb argument structures in the early years of morphosyntactic development. It aims to address, the descriptive problem, i.e. the emergent syntactic and semantic features of the child's verb usage and the developmental trajectory of verb argument structure. It also tries to tackle the explanatory problem, i.e. to what extent do Mandarin-specific morph-syntactic properties affect the path of the acquisition of verb argument structure?

The results show three prominent features of the early use of verbs in the child speech in Mandarin: dominance of transitive verbs, frequent argument omission, and theme-oriented arguments. Sentences with only one argument emerge first and remain the most productive across the early stages of syntactic development. Sentences without any overt argument are also quite frequent at the early one-word stage. Similar pattern of argument omission was reported in Huang's (2012) analysis of two Mandarin-learning children's speech between ages 2;2 to 3;1. She found that very small proportions of the children's transitive verbs (2-3%) contained two lexical arguments. Most of the transitive verbs contain only one or null lexical arguments and intransitive verbs tend to have null arguments.

The realization of overt arguments does not seem to result from limited processing abilities of the child at the very young age as revealed by that (1) objects of transitive verbs are more likely to maintained than subjects; (2) subjects of intransitive

verbs are more likely to be maintained than being dropped; and (3) even at the earliest developmental stage (i.e. the mean MLU close to 2 words), about 20% of the transitive verbs have a full realization of two arguments (cf. Figure 2). Overall an argument is likely to occur either in a subject or an object position, but more likely in object positions for transitive verbs and subject positions for intransitive verbs. The acquisition of arguments proceeds in the order of theme, agent, location, and recipient. Theme-only arguments emerge early and are the most productive. Agent-only and agent-theme structures increase constantly with the increase of the MLU.

A number of factors may have contributed to the acquisition pattern observed above, including certain cognitive bias that children bring to verb acquisition, adult production of verb argument structure in input, and discourse pragmatic features of Mandarin. The early emergence and dominance of transitive verbs may be related to children's sensitivity to the so-called "manipulative activity scene" (Slobin 1985), a causal event in which an agent intentionally brings about a physical and perceptual change of state or location on a patient by means of direct body contact or instrument. The high saliency of such events may draw children's attention and presents a good start for the child to map agent-patient relations onto language-specific forms.²

The distributional features of adult use of verb arguments may be a confounding factor for children's frequent production of transitive verbs and 1-argument-only structures. For example, transitive clauses were favored in Chinese both adult-to-adult conversations and written discourses (Lin 2009). Tao and Thompson (1994) showed that the majority of transitive clauses contain only one overt argument and the non-transitives tend to sustain the single lexical argument. Liu (1995) also observed that half of the intransitives and transitives had only one overt argument in spoken and written discourse in Mandarin. Lee and Naigles (2005) examined 7,884 tokens of the 60 most frequent verb types from caregiver speech to 10 different Mandarin-learning children aged around 1;10. They found a probabilistic association between number of arguments and verbs of different semantic classes for Mandarin, e.g. object NPs followed transitive verbs significantly more frequently than intransitive verbs.

Discourse pragmatic factor may also contribute to the argument realization in terms of information structure, especially the Preferred Argument Structure (Du Bois 1987). The Preferred Argument Structure argues that no more than one lexical argument is preferred; that lexical argument does not appear in the Agent role (subject of a transitive verb); that each clause contains no more than one argument carrying new information; that new information is introduced into discourse through O (object of a transitive verb) or S (subject of an intransitive verb). Given information is usually verbalized with pronouns or ellipsis, whereas new information typically realized with lexical nouns or phrases. In the current data, the majority of the transitive verbs tend to

---

² Note that this perceptual bias may not function with the same weight for verb learning across languages. For example, Ryan (2008) found that intransitives (unaccusatives and anticausatives) were the first to appear in a Spanish-learning child's verb development and transitive verbs appeared after intransitives.

have an object argument and intransitive verbs tend to maintain one subject argument. This pattern corroborates previous findings that children follow the Preferred Argument Structure (see review in Allen 2009).

The current study reveals prominent linguistic features in a Mandarin-learning child's production of verb argument structure. The acquisition trajectory suggests the child's early sensitivity to argument ellipsis in Mandarin and the adult production of verb argument structure. It reflects the effect of some possibly language-general cognitive mechanism and the discourse pragmatic properties of Mandarin. This study poses a first step in understanding the emergence and development of verb argument structure in Mandarin. This study concludes with more questions than answers as many questions remain to be explored. For example, can we generalize the current acquisition pattern of one child to other Mandarin-learning children? How does adult production of verb argument structure in child-directed speech influence the acquisition trajectory? What linguistic and contextual resources do children resort to for determining the full argument structure of the verbs? More crosslinguistic comparisons between typologically different languages are also needed in order to shed light on our understanding of the language-general and language-specific factors in language acquisition.

REFERENCES

- ALLEN, SHANLEY. 2000. A discourse-pragmatic explanation for argument representation in child Inuktitut. *Linguistics* 38.483-521.
- ALLEN, SHANLEY. 2009. Argument structure. *Cambridge Handbook of Child Language*, edited by E. Bavin, 217-36. Cambridge: Cambridge University Press.
- BROWN, ROGER. 1973. *A first language: The early stages* Cambridge, MA: Harvard University Press.
- CHEN, JIDONG. 2008. The acquisition of verb compounding in Mandarin. Nijmegen & Amsterdam, the NL: The Max-Planck Institute for Psycholinguistics & Free University of Amsterdam Doctoral dissertation.
- CHEUNG, HIN-TAT. 1998. Pingjun huayu changdu zai zhongwen de yingyong (The application of MLU measurement to Mandarin Chinese). *Ting Yu Hui Kan (Journal of Hearing Science)* 13.36-48.
- CHOMSKY, NOAM. 1999. On the nature, use, and acquisition of language. *Handbook of child language acquisition*, edited by W.C. Ritchie & T.K. Bhatia. Cambridge: Cambridge University Press.
- DU BOIS, JOHN W. 1987. The discourse basis of ergativity. *Language* 63.805-55.
- ERBAUGH, MARY. 1978. When minimal and simple don't mean the same thing... syntactic development of two Mandarin speaking two-year-olds.
- ERBAUGH, MARY. 1982. Coming to order: Natural selection and the origin of syntax in the Mandarin-speaking child. Berkeley, CA: University of California at Berkeley doctoral dissertation.

- ERBAUGH, MARY. 1992. The acquisition of Mandarin. *The crosslinguistic study of language acquisition*, edited by D.I. Slobin, 373-455. Hillsdale, NJ: Lawrence Erlbaum.
- HUANG, CHIUNG-CHIH. 2012. Preferred argument structure in Mandarin child language. *Taiwan Journal of Linguistics* 10.119-68.
- HYAMS, N. & K. WEXLER. 1993. On the grammatical basis of null subjects in child language. *Linguistics Inquiry* 24.421-259.
- LI, CHARLES & SANDRA THOMPSON. 1981. *Mandarin Chinese: A functional reference grammar* Los Angeles, CA: University of California Press.
- LIN, WAN-HUA. 2009. Preferred argument structure in Chinese: A comparison among conversations, narratives, and written texts. In *The Proceedings of the 21st North American Conference on Chinese Linguistics*. Rhode Island: Bryant Young University.
- LIU, SHU-MEI. 1995. Preferred clause structure in Mandarin spoken and written discourse. *Proceedings of the 10th Pacific Asia Conference on Language, Information and Computation*. 231-37.
- MACWHINNEY, BRIAN. 2000. *The CHILDES project: Tools for analyzing talk* Mahwah, NJ: Lawrence Erlbaum.
- MIN, RUI-FANG. 1994. The acquisition of referring expressions by young Chinese children: A longitudinal study of the forms and functions of early noun phrases. Nijmegen, the Netherlands: Catholic University of Nijmegen doctoral dissertation.
- RYAN, JOHN MICHAEL. 2008. L1 acquisition of argument structure in Spanish: Arizona State University doctoral dissertation.
- SLOBIN, DAN ISAAC. 1985. Crosslinguistic evidence for the Language-Making Capacity. *The crosslinguistic study of language acquisition*, edited by D.I. Slobin, 1157-256. Hillsdale, NJ: Lawrence Erlbaum.
- TAO, HONGYIN & SANDRA A. THOMPSON. 1994. The discourse and grammar interface: Preferred clause structure in Mandarin conversation. *Journal of Chinese Language Teachers Association* XXIX.1-34.
- TARDIF, TWILA. 1993. Adult-to-child speech and language acquisition in Mandarin Chinese. New Haven, CT: Yale University doctoral dissertation.
- TRASK, R. L. 1993. *A dictionary of grammatical terms in linguistics*. New York: Routledge.