Putting Object Before or After Verb: The Acquisition of VO and OV Word Order in Mandarin Chinese

Ya-Ching Yeh, University of Connecticut, USA
Letitia Naigles, University of Connecticut, USA

The European Conference on Language Learning 2014
Official Conference Proceedings

Abstract
Mandarin Chinese canonically uses SVO word order. However, a definite object can precede the verb, yielding OV order. While previous studies have established SVO order as Mandarin-speaking children's earliest and most frequent, little research has explored how they acquire OV order. The current study investigated this acquisition by analyzing Mandarin-speaking children's OV utterances.

We analyzed 40 children's spontaneous speech from the CHILDES Mandarin corpus; ten children in each of four age groups: 14-, 20-, 26-, and 32-month-olds. Qualified utterances (i.e., containing at least one verb) were coded for word order: **VO** (SVO, VO), **OV** (*Ba*OV, OV, OSV; SOV was excluded due to no occurrence), and Other. Frequencies were recast as percentages (of the total qualified utterances).
Compared to **VO** production, which appeared frequently from 20 months on, **OV** utterances emerged later. While both *Ba*OV and OSV have only sporadic occurrences in 26- and 32-month-olds, OV utterances grew significantly from 26 months on. By 32 months, children used OV at adult levels.

Further analyses reveal that the objects used in all OV utterances followed the definiteness constraint (e.g., definite nouns/demonstratives), and that the variety of verb types grew with age (i.e., extending from action verbs to cognition/desire/experience verbs). Appropriate verbal complements and/or aspect markers co-occurred with these verbs to make well-formed OV utterances. In sum, Mandarin-speaking children acquire the OV frame later than the VO frame, but they can produce OV utterances much like adults before age three, supporting the claim that grammatical competence is accomplished quickly in young children.

Introduction

Mandarin Chinese canonically uses SVO word order. SVO order is produced highly frequently in Mandarin, and the early sentences of Mandarin learners reflect this frequency, having been found to be consistently SV or VO under age two and strictly SVO between ages 2;0 and 2;9 (Cheng, 1986; Erbaugh, 1992).
In addition to acquiring the canonical word order, however, child Mandarin learners face a more complex situation, because Mandarin Chinese also allows a number of word order variants (Li & Thompson, 1981). Specifically, preverbal objects with and without *Ba* (to have), the definite/indefinite contrast, and ellipsis of noun arguments, all yield non-SVO orders. First, the *Ba* marker is used on objects in sentences in which the object precedes the verb (i.e., S*Ba*OV).

(1) wo3 ba3 wan3 xi3 le (S*Ba*OV)
I ba dish wash PFV (Perfective aspect) 'I washed the dishes.'

Second, while indefinite objects typically appear after the verb, definite objects can precede the subject and verb, forming SOV and OSV order.

(2) wo3 wan3 xi3 le (SOV)
I dish wash PFV (Perfective aspect) 'I washed the dishes.'

(3) wan3 wo3 xi3 le (OSV)
dish I wash PFV (Perfective aspect) 'I washed the dishes.'

Third, Mandarin allows ellipsis of noun arguments when they can be inferred from the context. Such ellipsis results in missing subjects and/or objects. Therefore, both SOV and OSV orders might become OV when the subject is omitted.

(4) wan3 xi3 le (OV)
dish wash PFV (Perfective aspect) '(Someone) washed the dish.'

While previous studies have established SVO order as child Mandarin learners' earliest and most frequent, little research has investigated how non-canonical word orders are acquired. The purpose of the current study is to investigate Mandarin learners' acquisition of non-SVO orders. The specific research questions addressed are as follows:

(1) How are Mandarin learners' non-SVO orders distributed as they grow?
We examined frequency distribution of non-SVO orders in child production and age-related changes that occurred in the acquisition process.

(2) Does Mandarin learners' non-SVO production follow the definiteness constraint and the complexity constraint?
The definiteness constraint means that the object in a non-SVO utterance has to be definite (Li & Thompson, 1983). The complexity constraint means that, if the bare form of a verb is not qualified for a full predicate in an utterance, the verb has to be morphologically complex, co-occurring with an appropriate verbal complement and/or an aspect marker (Du, 2006; Xu, 2012). By reviewing non-SVO utterances, we

examined whether Mandarin learners follow these two constraints while using non-SVO orders.

(3) How does the variety of verb types change in Mandarin learners' non-SVO production?
We also examined the verb types used in non-SVO utterances, which have not been investigated in previous studies (e.g., Cheng, 1986; Erbaugh, 1992), to see how the variety of verb types changes during the acquisition of non-SVO orders.

Method

Participants

The participants consisted of 40 Mandarin Chinese-speaking children selected from the CHILDES Mandarin corpus (MacWhinney, 2000; Zhou, 2000). All participants were native Mandarin speakers living in Nanjing, China. The children were from one of four age groups: 14-, 20-, 26-, and 32-month-olds. The numbers of girls and boys within each group were equal. All the children showed no hearing impairment or developmental delay.

Procedure

The mother-child conversation and interactions were video-recorded in each child's own day care classroom. The mother and her child began semi-structured play, in which they played and talked using the contents of four boxes. The four boxes respectively contained (1) a ball, (2) a popular toy, (3) paper and crayons, and (4) a picture book with stories in Chinese. All dyads were told to explore all the four boxes in about 10 minutes.

Transcription and Coding

The transcripts were obtained from the CHILDES Mandarin corpus (MacWhinney, 2000; Zhou, 2000). Children's speech was used for data analysis. Coding was restricted to utterances appearing in their spontaneous speech. Therefore, all those utterances appearing in memorized or routine phrases, such as social routines (e.g. thank you, bye-bye, and so forth), songs, poems, nursery rhymes, and story narratives were excluded from coding. All of the children's utterances that fit the above criteria and included at least one verb and one object were manually parsed and coded for word order and verb type, i.e., the verb used in an utterance. For example, the verb used in "John pushed Mary" is "push".

Word order coding categories included Canonical (SVO, VO), Non-canonical (*Ba*OV, OV, OSV, SOV), and Other. Frequencies were then turned into percentages (of the total qualified utterances). Verb type coding was conducted for the utterances in Non-canonical categories, i.e., *Ba*OV, OV, and OSV.

Results

The first question addressed in the current study is whether Mandarin learners acquire non-canonical orders on a purely frequency-related basis. The frequency distribution

of canonical and non-canonical orders in child production is presented in Figure 1. As expected, canonical order appeared frequently from 20 months on (14% & 24%). This supports the previous findings (i.e., Cheng, 1986; Erbaugh, 1992) in that children grasp VO order as soon as they start to combine words.

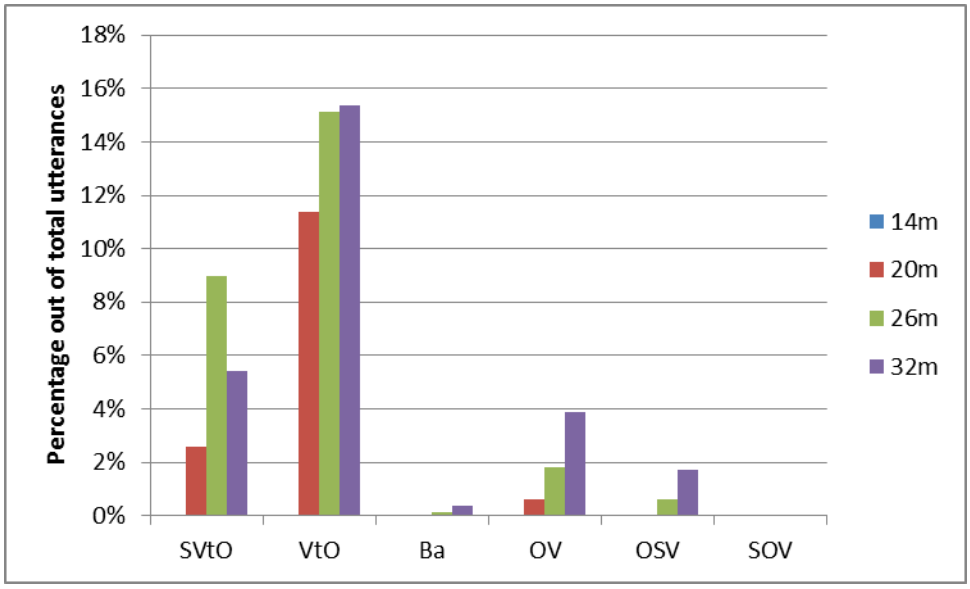

Figure 1. Frequency percentage of total qualified utterances of canonical and non-canonical word orders produced by the four age groups

Compared to their early SVO production, children produced non-canonical word orders much later and more rarely. As presented in Figure 1, the OV utterances began in 20 months and kept increasing through 32 months. There were sporadic occurrences with *Ba*OV and OSV orders only by 26- and 32-month-olds. No occurrences of SOV were observed at all; thus, SOV order was excluded from the subsequent analyses.

ANOVAs of the children's utterances (word order by age group) revealed a significant age group by order interaction, with only OV production showing significant age-related changes ($F(3, 36) = 4.449$, $p < 0.01$). Post hoc comparisons using the Tukey HSD test indicated that 32-month-olds (M = 3.88%, SD = 4.45%) produced significantly more OV utterances than 20-month-olds (M = 0.61%, SD = 1.33%) and 14-month-olds (M = 0%, SD = 0%). By 32 months, children used OV order at adult levels (3.88% vs. 3.73%; adult percentages were obtained from the same transcription, i.e., the speech of mothers who were videotaped with their children). Because only the OV utterances showed age-related changes, the subsequent analyses focused on the OV utterances.

The second question addressed in the current study first asks whether Mandarin learners follow the definiteness constraint, which requires a definite object to appear before the verb. As mentioned above, the OV utterances emerged in 20 months and increased significantly from 26 months on. The total number of OV utterances by these three age groups is 25: 20-month-olds produced 3 utterances while both 26- and 32-month-olds produced 11. Examination of these 25 OV utterances shows that all the objects appeared either as definite nouns or as demonstratives and pronouns. The definite nouns used are ones that appeared in ongoing conversation with mothers. The

demonstratives used include *zhe4ge*, 'this (one)' and *na4ge* 'that (one)'. Only one pronoun was used (by a 32-month-old). The objects and their frequencies appearing in these 25 OV utterances are presented in Table 1.

Table 1.
The object used in OV utterances

Age	Object used	English equivalent	Type
20m	da4de	The big (one)	definite noun
	zhe4ge (2)	This (one)	demonstrative
26m	zhe4ge (7)	This (one)	demonstrative
	na4ge	That (one)	demonstrative
	huai4dan4	The bad guy	definite noun
	yao1guai4	The devil	definite noun
	xiao3pang2yo3 de chi4che1	The car for kids	definite noun
32m	zhe4ge (6)	This (one)	demonstrative
	she2me	Whatever; Anything	pronoun
	xiao3chang2	The little bed	definite noun
	xiao3shiung2	The little bear	definite noun
	xiao3de	The small (one)	definite noun
	lan2se4	The blue (marker)	definite noun

Note: The numbers in parentheses indicate the frequencies of objects that appeared more than once

The second question also concerns whether Mandarin learners follow complexity constraint, which requires a verb to be morphologically complex, co-occurring with the appropriate verbal complement and/or an aspect marker when the bare form of the verb is not sufficient to act as a full predicate in an utterance.
In a Mandarin sentence, the predicate can be a bare verb. In Examples 5 and 6, the bare form of a verb is sufficient to describe its relation to the subject and/or object.

(5) wo3 zhi1dao4 zhe4ge ren2
I know this person "I know this person."

(6) zhe4ge bu2 yao4
this not want "(I) don't want this."

However, for some propositions/events, using only a bare verb may result in a semantically incomplete predicate (see Example 7).

(7) *zhe4ge fang4
this put "(someone) put this."

With a bare verb and no other grammatical units, Example 7 sounds unfinished and awkward. To make a semantically complete utterance, the verb *fang*4 (to put) in Example 7 must co-occur with other grammatical unites. According to the complexity constraint, possible grammatical units with which the verb *fang*4 (to put) can co-occur include various complements (e.g., adverb: *hao*3, "well"; *zai*4, "again"; locative complement: *zai*4*zuo*1*shang*4, "on the table") and aspect markers (e.g., perfective: '*le*'; durative: '*zhe*').

Of the 25 OV utterances, 21 utterances co-occurred with the verbal complements and/or aspect markers. Examination of these 21 utterances shows that appropriate verbal complements and/or aspect markers were used to make semantically complete and grammatically correct utterances. Examples are presented as follows.

(8) xiao3chang2 fang4 zhe4bian1 (verbal complement)
little bed put here "(I) put (the) little bed here."

(9) zhe4ge ba2 bu1 dong4 (verbal complement)
this pull not move "(I) can't pull this."

(10) zhe4ge yong4 kuo4 le (aspect markers)
this use EXP PFV "(I) have used this."

(11) da4de cha1 shang4 qu4 le (verbal complement and aspect marker)
big one insert up go PFV "(I already) inserted this (up there)."

EXP: experiential aspect PFV: perfective aspect

The other four utterances did not co-occur with verbal complements and/or aspect markers because the propositions of these four utterances did not require their verbs to follow complexity constraint. Instead, the verbs in these four utterances co-occurred with grammatical units, such as auxiliary, negation and V-not-V form. All these utterances are well-formed and acceptable. They are presented as follows.

(12) na4ge bu4nen2 wan2 (negation & auxiliary; by 26-month-old)
that not can play "(we) cannot play (with) that."

(13) zhe4 bu4 zhi1dao4 (negation; by 26-month-old)
this not know "(I) don't know this."

(14) she2me dou1 bu4 zhi1dao4 (negation; by 32-month-old)
what all not know "(I) don't know anything."

(15) zhe1ge yao4 bu2 yao4 (V-not-V form; by 32-month-old)
this want not want "(Do you) want this or not?"

The results from the above analyses reveal that, when Mandarin learners started to use OV order, they did produce well-formed OV utterances. That is, the objects and verbs that they used follow the definiteness constraint and the complexity constraint.
The third question addressed in the current study is how the variety of verb types changes when Mandarin learners acquire different non-canonical word orders. First, all the verbs used in the OV utterances were divided into two types: Action and Non-action verbs. Action verbs refer to verbs that describe action (e.g., walk) while non-action verbs describe state/situations that do not involve action (e.g., love) (Ross & Ma, 2006). The result is presented in Table 2. Both action and non-action verbs increased with age. Action verbs increased from 26 months on while non-action verbs increased from 32 months on.

Table 2.
Action vs. non-action verbs produced by 20-, 26- and 32-month-olds

Age	Action verbs	Non-action verbs
20m	cha1, '*insert*'	
	bai3, '*place*'	
26m	cha1, '*insert*'	gei3, '*give*'
	fang4, '*put*'	
	wan2, '*play*'	
	ba2, '*pull*'	
	da3, '*hit*'	
32m	fang4, '*put*'	gei3, '*give*'
	kai1, '*open*'	zhi1dao4, '*know*'
		yao4, '*want*'
		yong4, '*use*'

Second, all these verbs were further categorized into semantic classes. Six semantic classes were needed to classify 11 verb types found in these OV utterances. The result is presented in Table 3. All the verbs used by 20-month-olds were from the class of Manipulation. The classes of Motion and Change of Possession were added for 26-month-olds. 32-month-olds extended their verb use to the classes of Cognition, Desire, and Experience. In sum, more sematic classes appeared as children developed, suggesting that their variety of verb types grew with age.

Table 3.
Semantic classes used by 20-, 26-, and 32-month-olds

Age	Semantic class	Verbs
20m	Manipulation	cha1, '*insert*'; bai3, '*place*'
26m	Manipulation	cha1, '*insert*'; fang4, '*put*'; wan2, '*play*'
	Motion	ba2, '*pull*'; da3, '*hit*'
	Change of possession	gei3, '*give*'
32m	Manipulation	fang4, '*put*'
	Motion	kai1, '*open*'
	Change of possession	gei3, '*give*'
	Cognition	zhi1dao4, '*know*'
	Desire	yao4, '*want*'
	Experience	yong4, '*use*'

Discussion

The purpose of this study was to investigate the acquisition of non-canonical word orders in Mandarin Chinese by analyzing spontaneous speech of Mandarin learners of 14, 20, 26, and 32 months of age. The three research questions are addressed: (1) how Mandarin learners' non-SVO orders are distributed as they grow, (2) whether their non-SVO production follows the definiteness constraint and the complexity constraint, and (3) how the variety of verb types changes in their non-SVO utterances. The current study revealed that non-SVO word orders were produced later, less frequently than SVO orders by child Mandarin learners. This finding supports the

previous research (i.e., Cheng, 1986; Erbaugh, 1992) in that Mandarin learners produce VO order quite early when their two-word utterances emerge. This study also found that different non-SVO orders appeared with different frequencies. *Ba*OV, OSV, and SOV orders were rarely produced by Mandarin learners under age 3 while OV order increased significantly from 26 months on and reached adult levels before age 3. These OV utterances followed the definiteness constraint, carrying appropriate verbal complements and/or aspect markers, and showed a growing variety of verb types. These findings suggest that the acquisition of OV order is accomplished quickly in young children.

The current study found that OV order appeared earlier and more frequently than other non-canonical orders. Why did this happen? Frequency can be a factor. To investigate whether Mandarin learners acquire non-canonical orders on a purely frequency-related basis, further analyses on input will be needed.

In addition to the role of frequency, there might other possible factors that involved in the acquisition and resulted in the earlier production of OV order. The first possibility could be the meaning of *Ba* marker. Since the meaning of *Ba* marker is vague for young Mandarin learners, it is possible that they have not yet grasped what *Ba* marker means; thus, they did not use *Ba* construction confidently and frequently (Cheung, 1992). On the other hand, OV frame can be the result that the *Ba* marker is dropped in informal conversations for more concise communication (i.e., OV = (*Ba*)OV). Compared to the *Ba* construction, OV order is relatively simple and direct. As a result, using pure OV order would make production easier and more effortless for young children and further facilitate their OV occurrences. To test this possibility, we reviewed all the OV utterances produced by the children and their mothers (from the same transcription), and found that 19%-38% of the mothers' OV utterances have the potential to become *Ba*OV utterances. This suggests that Mandarin learners may hear OV utterances for *Ba*OV utterances with the *Ba* marker dropped. Our review of the children's OV utterances revealed that more than half of these utterances (55%-67%) have the potential to become *Ba*OV utterances. This suggests that Mandarin learners might choose OV order when they intend to use *Ba*OV.

The second possibility could be argument ellipsis, a language-specific property in Mandarin. Due to argument ellipsis, Mandarin speaker can drop subjects and/or objects from an utterance if either the subject and/or object can be inferred from the context and so does not need to be expressed overtly. If a conversation only involves the interlocutor as the subject in utterances, then this subject (i.e., either "I" or "you") is very likely to be omitted because the referent of the subject is obvious for both interlocutors. Thus, in a mother-child conversation, subjects such as "I", "you", and "we" would be likely to be dropped by the interlocutors. When an object acts as a topic and introduces new information, then it cannot be dropped in most cases. With these ellipsis properties, Mandarin learners may actually use OV order for their OSV and SOV utterances. This may explain why we found the higher frequency of OV utterances than those of OSV and SOV utterances in Mandarin learners' production.

The current study used the CHILDES Mandarin corpus (MacWhinney, 2000; Zhou, 2000), with the transcription from 10-minute mother-child interaction. The data generated from such a short period could be limited in information that can reveal children's language competence. Thus, it could be that Mandarin learners make more

*Ba*OV, OSV, and SOV utterances than what we found in the current study. To test this possibility, further investigation of other corpus will be needed.

Taken together, we found that Mandarin-speaking children's acquisition of non-SVO word orders occur later compared to their acquisition of SVO orders. Although the total number of non-SVO utterances is much smaller than that of SVO utterances, Mandarin learners still demonstrate well-formed OV utterances, suggesting that grammatical competence is accomplished quickly by young children.

References

Cheng, S. (1986). The earliest grammatical rules in Chinese toddlers and their probable developing process. *Chinese Journal of Psychology, 28*, 93-122.

Cheung, H. (1992). The acquisition of BA in Mandarin. Doctoral Dissertation, University of Kansas.

Du, H. (2004). The acquisition of the Chinese BA-construction by adult second language learners. Doctoral Dissertation, University of Arizona.

Erbaugh, M. S. (1992). The acquisition of Mandarin. In D. Slobin (Ed.), *The crosslinguistic study of language acquisition: Vol 3*. (pp. 373-455). Hillsdale, NJ: Erlbaum.

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

MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk* (3rd ed.). Mahwah, NJ: Erlbaum.

Ross, C. & Ma, J. (2006). *Modern Mandarin Chinese grammar: A practical guide*. New York, NY: Routledge.

Zu, H. (2012). *The acquisition of the BA construction by English-speaking learners of Chinese*. Doctoral Dissertation, University of Kansas.

Zhou, J. (2000). *Pragmatic development of Mandarin-speaking young children from 14 months to 32 months*. Doctoral Dissertation, The University of Hong Kong.

Contact email: myeh16@gmail.com