Syntactic Complexity and Productivity: A Study of Early Verbs in L1 Acquisition of Mandarin Chinese*

Xiaolu Yang
Tsinghua University, Beijing

An issue in the study of language acquisition that has attracted much attention is the nature of early syntax. At around 2, children start to combine words. Such multi-word combinations are characteristic of telegraphic speech. The nativist approaches to language acquisition believe that these early multi-word combinations are quite complex, with basic grammar of the ambient language already in order, including those structures whose subtle and sophisticated properties are not easily detected in the linguistic input. It is suggested that language acquisition is guided by innate linguistic mechanisms, in the form of Universal Grammar (Chomsky 1981; also see Pinker 1984, 1989, Crain 1991, Wexler 1998). Children's syntax is rule-governed and develops fast and catastrophically. Following this view, young children have access to syntactic categories, including both lexical categories such as nouns and verbs and functional categories from early on. In other words, grammatical categories are expected to be operative when the child starts to produce sentences. Thus, the superficially simple and sometimes erroneous sentences in early language development are in fact constrained by linguistic principles. The usage-based accounts of language acquisition, on the other hand, challenge the nativist view (Tomasello 1992, 2000, 2003; see also Lieven et al 1997, Pine and Lieven 1993, Pine et al 1998). The major tenet of such accounts is that early syntactic development is characteristic of item-based learning and takes place gradually and in piecemeal fashion, driven by the input data the child is exposed to. Language acquisition is supposed to be guided by general cognitive principles rather than innate language-specific principles such as UG.

The present study is an attempt to explore the nature of early syntax by looking at early verbs in Mandarin Chinese. The next section is a brief review of the usage-based view of early verbs. We will then present our findings. Discussion and conclusion will come in the last section.

1. The usage-based view of early verbs: the Verb Island Hypothesis

Tomasello (1992) claims that early verbs develop along different paths. Verbs operate as "individual islands of organization." (1992:257). Young children do not yet possess verbs as a general abstract grammatical category, but rather they are picking up verbs one by one individually. Children are conservative learners: they construct new combinations out of previously used materials, only in the way they hear them used. So, young children's productivity with the use of verbs is limited and newly learned structures seldom 'transfer across verbs.' Early multi-word combinations are simple, mostly derived on the basis of general cognitive processes such as symbolic integration (Tomasello 1992; also termed 'structure combining' in Tomasello 2000a). Major evidence in support of the Verb Island Hypothesis comes from Tomasello's diary study of early verbs by an English-speaking child, namely T, whose use of verbs was observed and recorded from 16 months to 24 months. It is found that T constructed her more complicated structures out of the materials that were used previously in

* The present study is part of the Chinese Early Language Acquisition (CELA) project, led by Prof. Thomas Lee，who was supported by a Hong Kong RGC grant（CityU 1245/02H). We thank CUI Rongjia, DENG Xiangjun，Ji Sunjing, Xiao Dan, Xu Yanqin and YANG Li for data collection and processing. This study is also supported by a national social sciences grant (05BYY038) to Xiaolu Yang.

less complex structures. Around “92% of T’s first 271 three-or-more-word sentences involved only a single simple change from previous sentences with that same verb.” (1992:236) Similar results are also obtained in Lieven *et al*. (1997) and Pine *et al*. (1998).

Following Ninio (1988), Tomasello (1992) suggests that early verbs should not form an abstract, coherent category if they do not appear as arguments of other predicates. Verbs that appear after modals verbs or in serial verb constructions, for instance, are arguments of modals and other predicates. Verbs with present or past tense morphology are also signs of verbs as arguments of higher predicates. Only when children show use of verbs in this way can one say that verbs are possessed as a syntactic category. In contrast to early nouns which frequently occurred as arguments of verbs, T in Tomasello’s (1992) study did not show any sign of using verbs in the argument slots of higher predicates: her verb morphology was inconsistent and newly learned verbs only entered into a limited range of sentence types and seldom appeared in the argument slots.

One question that immediately arises now is whether the Verb Island Hypothesis could account for early verbs in Mandarin Chinese: a language that is well-known for a lack of overt verb morphology. If what Tomasello and his colleagues have suggested so far is correct, we would expect similar results in L1 acquisition of verbs in Mandarin Chinese, that is, early use of verbs would be simple and limited, with different verbs behaving in different ways.

2. The present study

2.1 Methodology

In the present study, we used transcripts of two Mandarin-speaking kids, CY and ZHZ, from the CELA (Chinese Early Language Acquisition) project. Both kids were born and now live in Beijing. The two kids were visited weekly or biweekly for an extended period of time, one hour at each visit. The basic information of the two kids is summarized in Table 1.

Table 1. Information on two Beijing girls: CY and ZHZ

Name	CY	ZHZ
Period observed	0;10;04-1;11;30	1;04;19-1;11;17
Number of sessions used for analysis	53	24

We extracted and analyzed multi-word sentences containing verbs produced by CY and ZHZ during the observed period (0;10;04-1;11;30 for CY and 1;04;19-1;11;17 for ZHZ). If verbs in early Mandarin are of the same developmental pattern, just as what has been reported for L1 acquisition of English verbs, we would expect that: a) there is no sign of CY’s and ZHZ’s early verbs as a general and coherent word class; b) their sentences containing verbs in the observed periods should be relatively simple, constructed via simple symbolic integration processes such as coordination, addition, or expansion (Tomasello 1992); c) their early verbs should be limited in their use. i.e. occurring mostly in limited sentence patterns.

2.2 Major findings

Altogether CY produced 880 multi-word combinations containing verbs and ZHZ 693. During the observed period, one argument sentences and two argument sentences were common in the two kids’ transcripts. They were quite productive in expressing AGENT and THEME. Over 25% of CY’s early sentences contained

AGENT and 17% of ZHZ's did so. All AGENTs appeared in the subject position. THEME appeared in about 50% of CY's sentences and 40% of ZHZ's sentences. Arguments like EXPERIENCER, RECIPIENT and LOCATION also appeared at this stage, though their production was still limited. Table 2 is a summary of argument types produced by CY and ZHZ.

Table 2: Major argument types

Argument types / Age	CY	ZHZ
One argument sentences		
Agent (e.g. *ayi chi* aunt eat)	148 (16.8%)	84 (12%)
Theme (e.g. *na zhi* hold paper)	258 (29.3%)	210 (30%)
Recipient (e.g. *gei jiejie kan* to sister see)	16 (1.8%)	25 (3.6%)
Location (e.g. *tian-shang* fei sky-on fly)	26 (2.9%)	17 (2.5%)
Experiencer (e.g. *mama kan* mother see)	6 (0.7%)	8 (1.2%)
Two argument sentences		
Agent-theme (e.g *da hui lang chi ni* big grey wolf eat you)	83 (9.4%)	35 (5.1%)
Theme-location (e.g. *qingwa zai zhe* frog at here)	43 (4.9%)	36 (5.2%)
Experiencer-theme (e.g. *zanmen ting yinyue* we hear music)	9 (1%)	3 (0.4%)
Recipient-theme (e.g. *gei ta wei shui* give it feed water)	5 (0.6%)	2 (0.3%)
Total no. of multi-word sentences	880	693

Both kids tended to be very sensitive to the word order of Mandarin. There were very few cases of incorrect positioning of agents and themes at this stage. They consistently put the agent in the subject position of their one-argument and two argument sentences, as already mentioned. Themes mostly occurred in the pre-verbal position, as arguments of transitive verbs, mostly action verbs with a few exceptions of psychological verbs like *xihuan* ('like') or verbs of existence like *you* ('have'). But there were also cases when the subject of the sentence was theme. In such cases, verbs were either unaccusatives or unergatives. The position of themes thus invites one to think that CY and ZHZ were aware of different verb types. During this period, however, both kids produced a large number of sentences with no subjects nor objects. It is quite natural, as Chinese is a language allowing null subjects/objects (Li and Thompson 1981). Our children's drop of subjects or objects demonstrates their early sensitivity to this property of Mandarin Chinese.

ZHZ's marking of locative arguments has caught our attention. Tomasello (1992) noted that his subject T used prepositions to mark locative argument before 2, but her use was quite inconsistent, with some marked and others unmarked. ZHZ expressed location in 53 sentences. What's interesting is her use of the locative particle *zai* ('at') and *shang* (on). *Shang* in Mandarin usually follows the noun phrase that specifies a location. ZHZ used *shang* with 19 different noun phrase and this involved 10 different verbs. It seems that using *shang* to mark a locative argument was quite productive in ZHZ's early language, though CY did not use *shang* that much at this stage.

Verbs in Mandarin Chinese lack overt tense and agreement morphology, but one could easily determine the status of a verb in terms of its co-occurrence with other words. To see whether the two kids' early verbs formed a general and coherent class, we looked for the possibility of whether their verbs co-occurred with functional categories such as negators, modals, aspect markers or focus adverbs. The logic is that if CY and ZHZ consistently used such categories with verbs and place them in the right position (usually pre-verbally except aspect markers), and such a use was not restricted to a limited number of verbs, there is no reason not to believe that their verbs form a coherent word class. This would also serve as evidence that their early multi-word

combinations are not formed via simple symbolic integration processes such as coordination, addition, or expansion (Tomasello 1992), since sentences containing negators, modals, aspect markers and focus adverbs involve more complex processes. Table 3 displays the number of verbs that co-occurr with other categories in the transcripts. We could see that modals and aspect markers were not productive at all at this stage and could be ignored. The use of negators and focus adverbs, however, is quite revealing: both the kids used negators and focus adverbs in some sentences and these words were almost all placed correctly before verbs. Besides, negators and focus adverbs did not co-occur only with a limited number of verbs: 39 out of 135 verbs in CY's transcripts co-occurred with negators and 35 out of 162 verbs did so for ZHZ. When they started using the negator (i.e. the 1;5-1;7 period), they already used it with different verbs (with 6 different verbs for CY and with 7 different verbs for ZHZ) and the number of co-occurrences increased as they became older.

Table 3. Number of verbs co-occurring with various functional categories

Age / Structure	1;5-1;7	1;8-1;9	1;10-1;11	Total no. of verbs in the structure
CY				
With negators	6	9	32	39/135 (28.9%)
With focus adverbs	3	6	17	23/135 (17%)
With aspect markers	0	0	9	9/135 (6.7%)
With modals	1	3	6	10/135 (7.4%)
ZHZ				
With negators	7	15	20	35/162 (21.6%)
With focus adverbs	0	4	15	17/162 (10.5%)
With aspect markers	1	2	4	5/162 (3.1%)
With modals	0	1	2	3/162 (1.9%)

One important type of evidence in support of the Verb Island Hypothesis is that verbs are rarely found in the argument slots of higher predicates. Tomasello observes in his English-speaking subject T's data, "verbs did not appear in argument slots at all."(1992:252) What about our Mandarin-speaking subjects? To answer this question, we checked CY's and ZHZ's sentences containing two verbs. We calculated the second verb in such sentences. The results are given in Table 4. One could see that both produced serial verb constructions, particularly CY. About one third of verbs in CY's data entered into the serial verb construction. At the beginning stage, that is during the 1;5-1;7 period, the serial verb construction was not productive, but it turned to be so in the next two months (1;8-1;9).

Table 4: Number of verbs in the serial verb construction

Age / Subject	1;5-1;7	1;8-1;9	1;10-1;11	total no. of verbs
CY	2	12	28	36/135 (27%)
ZHZ	2	13	13	24/162 (15%)

In addition to serial verb constructions, our kids' early verbs also feature a complexity in the VP-internal structure, as evident in their V-V compounds, V-reduplication, and V-*yi*-*xia* structures. V-V compounding is a very important and very productive process in Mandarin Chinese. Its complexity lies in the interaction of semantics and syntax and RVCs (Resultative Verb Compounds) particularly raise a lot of significant issues

concerning the argument structure of the verbs involved and as a composite (Li 1990, Cheng and Huang 1994). Observe the example in (1a) produced by CY at 1;8. In the sentence, verbs *ti* (‘kick’) and *dao* (‘dao’) form a RVC *ti-dao*, the first verb denoting an action of ‘kicking’ that leads to the result of ‘falling down’ expressed by the second verb. The V-reduplication and V-yi-xia structure denote doing an action ‘a little bit’ or for a short period of time. The *yi* (‘one’) and *yi-xia* (‘one-down’) in V-*yi*-V and V-*yi-xia* function like a quantity adverbial that specifies the extent or duration of an activity. See the samples in (2-3).

(1) Gao lou ti-dao le. (CY, 1;8)
Tall building kick-fall LE.
‘The tall building was kicked down.’

(2) V-*yi-xia*
mama an yi-xia (CY, 1;10)
mother press one-down.
‘Mom presses (it) a little.’‘

(3) V-reduplication
a. mo-mo jiao. (CY, 1;06)
touch-touch foot
‘(You) touch the foot a little.’
b. wan-yi-wan (ZHZ, 1;8).
play-one-play.
‘(You) play a little.’

Table 5 displays the number of verbs that appear in these structures. We could see that before 2, V-V compounds and V-reduplication were not uncommon. Around 15% of different verbs appeared as the second verb in the compounds for both kids and a similar number of verbs also appeared in V-reduplications. Compared with V-V compounds and V-reduplication, V-*yi-xia* structures were still rare at this stage.

Table 5. Number of verbs in V-V compounds, V-reduplication and V-yi-xia structures

Age / structures	Before 1;5	1;5-1;7	1;8-1;9	1;10-1;11	Total
CY					
V-V compounds	1	8	11	18	20/135
V-reduplication	9	6	4	8	20/135
V-*yi-xia*	1	2	2	1	6/135
ZHZ					
V-V compounds	0	3	16	16	24/165
V-reduplication	0	2	18	10	28/165
V-*yi-xia*	0	0	1	5	6/165

Though English-speaking children are reported to be quite limited in the use of almost all their early verbs (Tomasello 1992, Lieven et al 1997), i.e. using them only in one sentence pattern (our ‘sentence pattern’ is defined as any change in argument types, functional categories, internal VP structure, etc.), the picture with Mandarin-speaking children appears to be a different one. What Table 6 tells us is that early verbs in the two Mandarin-speaking children’s production were not limited to only one sentence pattern. CY was particularly

productive in her early use of verbs. Except the period before 1;5, the number of her verbs that entered into two or more sentence patterns was similar to or even more than that of verbs employing one sentence pattern (22 vs 31 in the 1;5-1;7 period, 38 vs. 30 in the 1;8-1:9 period and 62 vs. 41 in the 1;10-1;11 period). ZHZ's data also indicate that there was no lack of verbs that entered into two or more sentence patterns (38 in the 1;8-19 period and 57 in the 1;10-1;11 period) before 2.

Table 6. Number of verbs entering into one and two or more sentence patterns

Age / Pattern	Before 1;5	1;5-1;7	1;8-1;9	1;10-1;11
CY				
One pattern	16	31	30	41
Two or more patterns	6	22	38	62
ZHZ				
One pattern	0	25	49	63
Two or more patterns	0	7	38	57

3. Conclusion

A current debate in the study of language acquisition pertains to whether early syntactic development is rule-based or usage-based. According to the proponents of usage-based accounts such as Tomasello (1992, 2000a, b), verbs in early child language develop along different paths, limited in use and simple in structure. In the present study, we analyzed multi-word combinations containing verbs before 2 by two Mandarin-speaking children, CY and ZHZ. We looked in particular at several structures: the use of functional categories such as negators and modals in such combinations and the use of two verbs. Contrary to Tomasello's (1992) findings that his English-speaking subject's early verbs are characteristic of "concreteness, particularity, and idiosyncracy" (p.264), our data demonstrate some generality and systematicity in young Mandarin-speaking children's use of early verbs. We have seen that quite a number of different verbs appeared as arguments of higher predicates: negators and modals consistently appeared pre-verbally. Many different verbs appeared in serial verb constructions and in V-V compounds. It was also common for the two kids to use verbs in two or more patterns during the observed period. ZHZ's use of locative particle *shang* ('on') suggests a productive process of locative argument marking. We would take such findings as evidence that early verbs in Mandarin do form a general and coherent syntactic category and that early child Mandarin is quite productive and complex.

Our findings are consistent with findings of some recent studies of acquisition of Chinese, which have posed problems for the usage-based accounts of early syntax. Lee (2004) has presented data from two Cantonese children (one observed from 1;5 to 2;7 and the other from 1;6 to 2;8). The data do not provide evidence for the lexically-based account of early child speech as proposed in Lieven et al (1992, 1993, 1997). Zhang et al's (2005) study of sentence-final particles (SFPs) used by young Mandarin-speaking children (before 3) suggests that SFPs in early Mandarin development form a syntactic category rather than specific lexical items. The emergence of SFPs reflects a rapid, catastrophic process rather than a gradual, piecemeal one. Yang and Xiao (2005) explore the development of the BA construction of a two-year-old Mandarin-speaking child and find that it

cannot be accounted for by imitative learning nor cognitive symbolic integration processes as proposed in Tomasello (1992). Why should Chinese-speaking children's early syntactic development differ from that of English-speaking children as reported in studies by Tomasell and his colleagues? One might attribute this to a difference in languages and to the different linguistic environment that children are exposed to. If we do so, then we would be faced with a serious problem in accounting for the fact that children with different L1 backgrounds will undergo similar developmental stages such as the drop of null subjects/objects (Hyams 1986, 1996) and the root infinitive stage (Hoekstra and Hyams 1998, Wexler 1998). We would also find it difficult to explain English-speaking children's non-adult wh-questions with an extra 'media' wh-word (Thornton 1990).

Our findings reflect our children's early sensitivity to many aspects of basic grammar of Mandarin Chinese, including word order, null subjects/objects, V-V compounds, etc.. This early sensitivity might be better explained as a result of very early parameter setting in Wexler's sense (1998). The productive use of verbs in different patterns and the complexity of VP structures in early child Mandarin Chinese probably may have to do with children's early knowledge of some abstract verbs such as DO, EXIST, CAUSE, BECOME, BE, etc. (Hale and Keyser 1990, Huang 1997, Lin 2001). Whether this is the case is indeed worth further longitudinal and experimental investigation.

References

Crain, S. 1991. Language acquisition in the absence of experience. *Behavioral and Brain Sciences* 14:597-650.

Hale, K. and Keyser, S.J. 1991. On the syntax of argument structure. Lexicon Project Working Papers, Center for Cognitive Science, MIT.

Hoekstra, T. and Hyams, N. 1998. Aspects of root infinitives. Lingua

Huang, C.T.J. 1997. On lexical structure and syntactic projection. Symposium Series of the Institute of History and Philology, Academia Sinica, Number 2.

Hyams, N. 1986. *Language Acquisition and the Theory of Parameters*. Dordrecht: Reidel.

---1996. The underspecification of functional categories in early grammar. In H. Clahsen (ed.), *Generative Perspectives on Language Acquisition*. Amsterdam: Benjamins.

Lee，T. 2004. Productivity in the early word combinations of Cantonese-speaking children. In Shi Feng and Shen Zhongwei (eds.), *The Joy of Research – A Festschrift in Honor of Professor William S.Y. Wang on His Seventieth Birthday*. Tianjin: Nankai University Press.

Li, C.N. and Thompson, S.A. 1981. *Mandarin Chinese: A Functional Reference Grammar*. Berkeley, CA.: University of California Press.

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

Lin, T.H. 2001. Light Verb Syntax and the Theory of Phrase Structure. Ph.D. dissertation, University of California, Irvine.

Nino, A. 1988. On formal grammaticall categories in early child language. In Y. Levy, I. Schlesinger, and M. Braine (eds.), *Categories and Processes in Language Acquisition*. Hillsdale, NJ.: Erlbaum.

Pine, J., Lieven, E. and Rowland, G. 1998. Comparing different models of the development of the English verb category. *Linguistics* 36:4-40.

Pinker, S. 1984. *Language Learnability and Language Development*. Cambridge, Mass.: Harvard University Press.

--- 1989. *Language Learnability and Cognition: The Acquisition of Argument Structure.* Cambridge, MA: MIT Press.

Thornton, R. (1990) *Adventures in long-distance moving: the acquisition of wh-questions*. Doctoral dissertation. University of Connecticut.

Tomasello, M. 1992. *First Verbs: A Case Study of Early Lexical Development*. Cambridge: CUP.

--- 2000a. Do young children have adult syntactic competence? *Cognition* 74: 209-253.

--- 2000b. The item-based nature of children's early syntactic development *Trends in Cognitive Sciences* 4 (4)-153-163.

--- 2003. *Constructing a Language: a Usage-based Theory of Language Acquisition*. Harvard University Press.

Wexler, K. 1998. Very early parameter-setting and the unique checking constraint: a new explanation of the optional infinitive stage. *Lingua* 106:23-79.

Yang, X. and Xiao, D. 2005. The BA construction in early syntactic development of Mandarin: A case study. Paper presented at NACCL17, June 24-26, The Monterey Institute of International Studies, California, U.S.A.

Zhang, J. Fang, L. Lee, T. And. Yang, X. 2005. The acquisition of sentence final particles in Beijing Mandarin. Paper presented at IACL13, June 9-11, Leiden University.