Yu and Tamaoka *Lingua Sinica* (2018) 4:3
https://doi.org/10.1186/s40655-018-0035-x

Lingua Sinica

ORIGINAL ARTICLE Open Access

Age-related differences in the acceptability of non-canonical word orders in Mandarin Chinese

Shaoyun Yu* and Katsuo Tamaoka

* Correspondence: s.yu@nagoya-u.jp
Graduate School of Humanities, Nagoya University, Nagoya, Japan

Abstract

Prior theoretical and experimental work has demonstrated that the subject-object-verb (SOV) word order is more restricted and more difficult to process than the object-subject-verb (OSV) word order in Chinese. However, few studies have investigated the actual acceptability of non-canonical, verb-final word orders among native speakers. In this study, we conducted a paper-based survey with a group of younger adults and a group of older adults. Non-canonical sentences with and without animacy contrasts were tested with three tasks: acceptability judgement, grammaticality judgement, and subject selection. Our results have two main implications. First, contrary to previous studies, the results suggested that animate-animate-verb (AAV) sentences generally have very poor acceptability, and they may be mostly regarded as uninterpretable. Second, and importantly, we found that while a preference for OSV over SOV was observed in younger adults, it was not evident in older adults. This suggests that the preference for the OSV word order might be modulated by the factor of age.

Keywords: Word order, Animacy, Age, Acceptability, Grammaticality

1 Introduction

Chinese is canonically an SVO language.[1] Sun and Givón's (1985) survey of contemporary written and spoken Mandarin Chinese reports that over 90% of the direct objects occurred in the canonical position after the verb. At the same time, the non-canonical SOV and OSV word orders, with bare objects being placed in the sentence-medial or sentence-initial positions, are also possible in Chinese. For instance, 2a and 2b are the SOV and OSV variants of 1, with their propositional meanings being identical to 1.

(1) 张三擦了黑板
zhāngsān__cā__le__heībǎn
zhangsan__wipe__PERF__blackboard (SVO)
Zhangsan wiped the blackboard

(2) a. 张三黑板擦了
zhāngsān__ heībǎn__ cā__le (SOV)
zhangsan__blackboard__wipe__PERF

b. 黑板张三擦了
heībǎn__ zhāngsān__cā__le (OSV)
blackboard__zhangsan__wipe__PERF

While the OSV word order is generally considered to be the result of topicalization of the object (Huang et al. 2009), the syntactic status and function of the SOV word order are frequently debated. The prominent view holds that the object in the SOV word order is the 'focus', and it is obligatorily associated with a contrastive or emphatic reading (Ernst and Wang 1995; Shyu 1995, 2001). Paul (2002, 2005) challenged this view by demonstrating the distinctness of the sentence-medial object and the pre-verbal *lián* 'even' focus. Thereupon, Paul argues for a sentence-internal topic status for the sentence-medial object, whose function, similar to the sentence-initial topic, is to 'set the frame within which the main predication holds' (Paul 2015; Chafe 1976). More recently, Hsu (2008) has proposed that both the topic and the focus status are able to be assigned to the sentence-medial object.

Despite these debates, previous studies in theoretical linguistics have agreed that SOV is more restricted than OSV in Chinese. In particular, two syntactic constraints are known to be imposed on SOV but not OSV. First, as exemplified in the following sentences, while resumptive pronouns in the post-verbal position are allowed in OSV, they are prohibited in SOV (Huang et al. 2009).

(3) a. 张小姐$_i$,我不想追她$_i$
zhāng__xiǎojiě,__wǒ__bù__xiǎng__zhuī__tā (OSV)
zhang__Miss,__I__not__want__court__her
Miss Zhang, I do not want to court her.
b. * 我,张小姐$_i$不想追她$_i$
wǒ,__zhāng__xiǎojiě__bù__xiǎng__zhuī__tā (SOV)
I,__zhang__Miss__not__want__court__her
(Adapted from Huang et al. 2009)

Second, while an object inside an embedded clause is able to be extracted and relocated to the sentence-initial position, it is not allowed to be placed in the sentence-medial preverbal position (Lu 1994; Ernst and Wang 1995; Paul 2005):

(4) a. 王五说李四读完了**那本小说**
wángwǔ__shuō__lǐsì__dú__wán__le__**nà__běn__xiǎoshuō**
wangwu__say__lisi__read__finish__PERF__that__CL__novel
Wangwu said that Lisi finished reading that novel
b. **那本小说**王五说李四读完了
nà__běn__xiǎoshuō__wángwǔ__shuō__lǐsì__dú__wán__le
that__CL__novel__wangwu__say__lisi__read__finish__PERF
c. * 王五**那本小说**说李四读完了
wángwǔ__**nà__běn__xiǎoshuō**__shuō__lǐsì__dú__wán__le
wangwu__**that__CL__novel**__say__lisi__read__finish__PERF
(Lu 1994, from Ernst and Wang 1995)

SOV is also considered to be subject to additional semantic constraints in comparison to OSV. Hou (1979) noticed that the SOV word order is not available for animate objects. Sentences like 5c can only be interpreted as OSV, with a reversed reading of the original SVO one.

(5) a. 他批评了**那个女孩**
tā__pīpíng__le__**nèi__ge__nǚhái**
he__criticize__PERF__**that__CL__girl**
He criticized that girl
b. **那个女孩**他批评了
nèi__ge__nǚhái__tā__pīpíng__le
that__CL__girl__he__criticize__PERF
c. * 他**那个女孩**批评了
tā__**nèi__ge__nǚhái**__pīpíng__le
he__**that__CL__girl**__criticize__PERF
Intended reading: 'He criticized that girl'
(Adapted from Hou 1979)

However, it is crucial to point out that the subject of a sentence is also typically animate (Branigan et al. 2008), and the constraint on the SOV reading might arise from the lack of a contrast in the animacy of the arguments. Lu (1994) postulates that the object can only be fronted to the sentence-initial position when both the subject and object are animate. Qu (1994) further argues that when two preverbal nouns are equally likely to be the thematic agent, the non-canonical sentence is disambiguated as OSV, which is the 'default option'.

1.1 The processing of non-canonical word orders

While theoretical studies have examined the nature of non-canonical word orders with an emphasis on their syntactic properties and deviational constraints, studies in psycholinguistics and neurolinguistics have looked into the phenomena using experiments that investigate how these sentences are processed by comprehenders. When a comprehender is presented with a noun-noun-verb string, which of the two word orders (SOV or OSV) would be selected? Is there any difference in the processing of SOV and OSV?

The animacy of nouns has consistently been demonstrated to play a significant role in the interpretation of non-canonical sentences. Studies based on the Competition Model (Bates and MacWhinney 1981) have found that in Chinese, when a contrast of animacy in the preverbal nouns is displayed, there is an overwhelming tendency for the animate noun to be identified as the subject and the inanimate noun to be identified as the object (Miao 1981; Li et al. 1992; Li et al. 1993). These studies have measured the percentage of the first noun being chosen as the agent of the sentence; we will refer to this as the 'NP1 percentage'. While the NP1 percentage indicates the likelihood of the SOV interpretation, the complement of the NP1 percentage (i.e., the percentage of the *second* noun being chosen as the agent) indicates the likelihood of the OSV interpretation. In Miao (1981), animate-inanimate-verb (AIV) strings were interpreted as SOV 85.4% of the time, and inanimate-animate-verb (IAV) strings were interpreted as OSV

85.4% of the time too. In the offline experiment of Li et al. (1992), 96.5% of the AIV strings were interpreted as SOV, and 95.1% of the IAV strings were interpreted as OSV.

While the configuration of animacy (i.e., AIV or IAV) powerfully shifts the interpretation of non-canonical sentences, it is not the only force. In parallel to the theoretical studies that argued that SOV is more restricted than OSV, online sentence processing studies have suggested that SOV is also more difficult to process than OSV. An online experiment was reported in Li et al. (1992) as well as in Li et al. (1993). In this experiment, participants were given a maximum of 3s to choose the agent of a sentence; the reaction time for choices was measured in addition to the NP1 percentage. A slight asymmetry in the interpretation of AIV and IAV strings was discovered in the results. According to the NP1 percentage reported in Li et al. (1993), 98% of the IAV strings were interpreted as OSV; however, only 84% of the AIV strings were interpreted as SOV. The results suggested that under time pressure, a SOV word order disadvantage might exist. This factor might have caused a small portion of the responses to AIV strings to favour the OSV interpretation, despite the fact that there is a mismatch of the typical argument animacy in the OSV-interpreted AIV strings. Furthermore, according to Li et al. (1992) and Li et al. (1993), on average, it takes a longer time to decide the subject of AIV strings; the mean reaction time for the AIV strings (2162 ms) was statistically significantly longer than that for the IAV strings (1958 ms). Recent neurolinguistic studies also confirmed that there is a structural preference for OSV. In the ERP (event-related potential) study reported in Wang et al. (2012) and Wang 王路明 (2015), AIV strings with SOV readings elicited statistically significantly greater anterior negativity when compared to IAV sentences with OSV readings in the 450–700 ms time window post-verb onset, indicating that the SOV sentences are more difficult to process.

Another situation is the case of sentences without animacy contrast. In an animate-animate-verb (AAV) string, ideally the animacy is uniform between the two nouns, therefore any bias toward either the SOV or OSV interpretation should not be accounted for by noun animacy. To our knowledge, the available data in the literature are somewhat inconsistent. Miao's (1981) offline experiment suggested a SOV preference for the interpretation of AAV, as the NP1 was chosen as the agent 68% of the time. In contrast, the offline experiment from Li et al. (1992) revealed an OSV preference, with the NP1 being chosen as the agent 33.3% of the time. The online experiment reported in Li et al. (1992) and Li et al. (1993) closely replicated the offline results of Li et al. (1992); 34% of the AAV strings were interpreted as SOV, while 66% of the AAV strings were interpreted as OSV. Although Hou (1979), Lu (1994), and Qu (1994) have made rather strong claims that OSV is the sole interpretation of AAV strings, the online and offline results of Li et al. (1992) and Li et al. (1993) are only partially compatible with their perspectives, let alone the counterexample from Miao (1981). Even if we accept the results of Li et al. (1992) and Li et al. (1993), and agree that OSV is the preferred interpretation for AAV, there is still a reasonable amount of responses that favoured the opposite interpretation of SOV. Moreover, according to the data from Li et al. (1993), AAV strings on average took the longest time for participants to choose the agent (2662 ms). Taken together, it seems that rather than having a 'default' option in their mind, people might have difficulties in identifying the subject from two animate nouns, thus causing a considerable uncertainty in the interpretation of AAV strings and a

prolonged decision time. And if the thematic and syntactic roles are not determinable in the first place, the acceptability of a sentence is also in question. However, this question was not answered in previous experimental studies, as acceptability was not their main focus.

1.2 Age as a factor in language use and comprehension

Individual differences are seldom taken into account in studies on non-canonical word orders in Chinese. Theoretical studies often overlook the possibility that the perceived acceptability of a non-canonical word order might be systematically affected by certain individual factors such as age and gender. A majority of experimental studies contain only a single participant group (typically, young college students), and variables relevant to individual differences are usually not presented.

Age can have a significant effect on a speaker's use of language at various levels, such as pronunciation, word choice, and syntactic preferences (Coulmas 2005; Wagner 2012). As discussed by Meisel et al. (2013), while the core grammar of an individual is highly stable after childhood, age-related changes can nevertheless occur within the structural optionality provided by the grammatical system. Many studies in sociolinguistics have shown that speakers from different age groups can manifest quantitively different uses of alternative syntactic structures. For example, French exhibits two major variants to express future temporal reference, the inflected future (IF) and the periphrastic future (PF). Wagner and Sankoff (2011) analysed the use of IF and PF among Canadian French speakers and found that the conservative IF form is more frequently used by older speakers than by younger speakers. Moreover, they found that both groups of speakers have increased their use of IF as they age, suggesting an age-grading pattern in the use of the two alternative forms.

Age also has profound effects on human cognitive abilities (Craik and Bialystok 2006). In general, older adults are considered to have relatively worse performance in online sentence processing (Wlotko et al. 2010). Previous studies have demonstrated that older adults are likely to have more difficulties in comprehending structures that involve non-canonical word orders. In English, Stine-Morrow et al. (2000) found that older adults performed worse than younger adults in comprehension tasks for object relative clauses, while no significant age-related differences were observed for subject relative clauses. Caplan et al. (2011) also found that older adults' processing performance of object clefts, in addition to object relative clauses, was worse than that of younger adults. In Korean, a SOV language with relatively free word order, there is also evidence showing that older adults' performance in the processing of the non-canonical OSV word order is worse than that of younger adults (Oh et al. 2016; Sung 2017).

Developmental studies have revealed that both the use of non-canonical word orders and the preference for OSV over SOV in Chinese are not established until a rather late stage in the L1 acquisition of Chinese speaking children, which is in sharp contrast with their early and solid acquisition of the canonical SVO word order (Miao et al. 1984; Chen 2009; Lee 2010; Li 2010). Considering the possible effects of age on language use and sentence comprehension processing, it is natural to question whether the relative acceptability of the SOV and OSV word orders stays invariant throughout one's entire adulthood.

1.3 The present study

In this preliminary study, we conducted a paper-based survey focused on the acceptability of the non-canonical word orders in Chinese. In addition to the non-canonical AIV, IAV, and AAV strings, canonical SVO sentences were also included in the survey to serve as the baseline. We recruited a group of younger adults as well as a group of older adults to participate in the survey.

While studies based on the Competition Model have extensively studied the identification of thematic roles in non-canonical strings, the task of deciding one of the two nouns as the agent does not require a string of words to be fully acceptable. Although an acceptability judgement task was included in the ERP study of Wang et al. (2012) and Wang 王路明 (2015), it was a dichotomous yes-or-no task, and AAV strings were not among the objects of their research. To obtain a more fine-grained evaluation of the acceptability of the non-canonical word orders, we decided to assess the acceptability with a 7-point Likert scale. Because neutral responses are allowed on a 7-point scale, a grammaticality judgement task that allows only dichotomous yes-or-no responses was also included as a secondary index of acceptability.

Since an acceptable sentence requires the successful analysis of the syntactic and thematic roles of the arguments, it is important to know how the stimuli are interpreted by survey participants. Therefore, we added a subject selection task in our survey. As shown in previous studies based on the Competition Model, considerable uncertainty exists in the interpretation of AAV strings. We postulate that there might not exist a default interpretation for non-canonical sentences when the animacy of the two nouns is perfectly uniform, and this in turn would cause difficulties in analysing the syntactic and thematic roles in AAV strings. To validate our hypothesis, we included a 'cannot decide' option in the subject selection task.

According to previous studies, we expected that AIV strings would be interpreted as SOV, and IAV strings would be interpreted as OSV, with little to no responses choosing 'cannot decide'. We expected, however, that the acceptability of SOV would be lower than that of OSV, as SOV is subject to more restrictions, and it has been shown to be more difficult to process than OSV. On the other hand, for AAV strings, we predicted that a considerable amount of responses would fall into the 'cannot decide' option, as a result of failing to analyse the syntactic and thematic roles. Consequently, we predicted that AAV strings would have the lowest acceptability rating.

Sociolinguistics studies have shown that age can play a significant role in modulating the use of alternative syntactic structures in a language. In addition, experimental studies have demonstrated that the processing of non-canonical word orders could be more challenging for older adults. As optional structures with non-canonical word orders in Chinese, SOV and OSV might be subject to age-related changes. However, little work has been done to investigate whether the use or perception of SOV and OSV could change with age. By comparing the acceptability in the younger adults and the older adults, we were able to detect the possible differences in their subjective perception of the well-formedness of these word orders.

2 Methods

2.1 Participants

A total of 58 native Chinese speakers participated in the survey of this study. They were divided into two groups: the younger native speakers group and the older native speakers group (Fig. 1).

The younger group was comprised of 30 monolingually raised young native speakers of Chinese. One participant was excluded because of evidence of low attentiveness (frequently choosing objects as subjects in canonical SVO sentences). As the result of this exclusion, 29 participants were included in the final analysis. Their mean age was 25.8 years old, ranging from 22 to 32 years old. At the time of the survey, all participants of the younger group were students at Nagoya University, Japan.

The older group consisted of 28 middle- and older-aged monolingual native speakers of Chinese. Five participants were excluded because of evidence of low attentiveness. One participant frequently chose objects as subjects in canonical SVO sentences. Two participants frequently marked grammatical SVO sentences as ungrammatical. Two participants responded with an unnatural sequence of answers showing the same repetitious pattern in their responses. After this careful examination, 23 participants were included in the final analysis. Their mean age was 56.6 years old, ranging from 49 to 65 years old. At the time of the survey, all participants of the older group were residents of Zhenjiang City, Jiangsu Province, China.

2.2 Materials

The survey contained 120 test sentences, among which 40 were in the canonical SVO word order and 80 were in non-canonical verb-final word orders. Each sentence was formed with two nouns and one verb. All the nouns were double-character words. All the verbs were single-character words suffixed with the perfective aspect marker -了 *-le*.

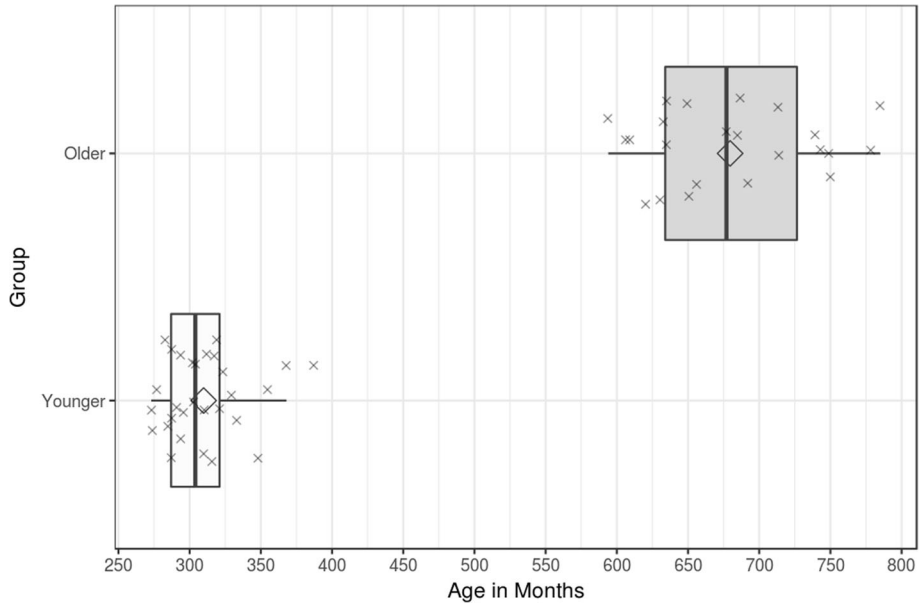

Fig. 1 Age distribution of the participants. Individual data points (represented by x-shaped symbols) and the mean values of each group (represented by rhombus-shaped symbols) are plotted over the boxplot

To generate the test sentences, we first prepared a list of 40 grammatical and natural sentences that were in the canonical SVO word order and were in the active voice. For 30 of these sentences, the subjects were animate nouns (i.e., human names) and the objects were inanimate nouns, e.g., 小王喝了牛奶 *xiǎo-wáng hē-le niúnǎi* 'xiao-wang[2] drank the milk'. The semantics of the verbs in these sentences does not allow the thematic roles of the two nouns to be exchanged. Since the two nouns differed in animacy, an *animacy contrast* was formed in such sentences. We denote this as [+animacy contrast]. For the other ten sentences in the list, both the subjects and the objects were human names, e.g., 小赵骗了小张 *xiǎo-zhào piàn-le xiǎo-zhāng* 'xiao-zhao deceived xiao-zhang'. As the two nouns were both human names, it is possible to exchange the thematic roles of the nouns. Since the two nouns were perfectly uniform in terms of animacy, there was no animacy contrast in such sentences. We denote this as [–animacy contrast].

Then, we derived a total of 80 non-canonical verb-final sentences from the 40 original SVO sentences. As exemplified in Table 1, for each original sentence, two derived sentences were generated by fronting the object before the verb. We denote a derived sentence as 'SOV' if the object was moved to the sentence-medial position immediately before the verb and as 'OSV' if the object was moved to the sentence-initial position. The terms 'SOV' and 'OSV' are surrounded with quotation marks because they do not necessarily represent how a verb-final sentence should be interpreted. Instead, they are merely intended to keep track of how the object from the original SVO sentence was positioned in a derived verb-final sentence. Different from Li et al. (1992) or Su (2001) in which test sentences were constructed randomly from pools of candidate nouns and verbs, the derivative approach used in the present study ensured that the verb-final sentences are all potentially acceptable, because it is assured that they could be rearranged as grammatical and natural SVO sentences.

The detailed composition of the 120 test sentences is shown in Table 2. The test sentences are cross-classified by two factors: *animacy contrast* and *word order*. As mentioned above, readers should be cautious that the *word order* factor in the present study only represents how the position of an object was manipulated during the sentence derivation process. For the [–animacy contrast] sentences, both 'SOV' and 'OSV' resulted in animate-animate-verb (AAV) strings, as the nouns were all human names without any bias. For the [+animacy contrast] sentences, however, 'SOV' resulted in animate-inanimate-verb (AIV) strings, and 'OSV' resulted in inanimate-animate-verb (IAV) strings. This is because in the original SVO sentences, the nouns for subjects were animate, and the nouns for objects were inanimate. As one of our main interests is to investigate how AIV and IAV sentences would differ in acceptability, the [+animacy contrast] sentences were given more weight in the survey. In the survey paper,

Table 1 Relationship between original SVO sentences and derived verb-final sentences

Animacy contrast	Word order		
	SVO (original)	'SOV'	'OSV'
+	小王 喝了 牛奶	小王 牛奶 喝了	牛奶 小王 喝了
–	小赵 骗了 小张	小赵 小张 骗了	小张 小赵 骗了

'SOV' and 'OSV' indicate how the object in an original SVO sentence is positioned in the derived verb-final sentence. Original objects are marked in bold typeface

Table 2 Detailed amount of the test sentences

Animacy contrast	Word order			Totals
	SVO (original)	'SOV'	'OSV'	
+	30	30	30	90
–	10	10	10	30
Totals	40	40	40	120

the order of the test sentences was randomised, and additional manual adjustment was made so that sentences in derivational relationships would not appear sequentially.

2.3 Procedure

For each test sentence in the survey paper, participants were asked to complete three tasks about the sentence in the following order: (1) acceptability judgement, (2) grammaticality judgement, and (3) subject selection. In the acceptability judgement task, participants were required to rate the acceptability of a test sentence on a 7-point Likert scale. In the grammaticality judgement task, participants were required to decide whether a test sentence was grammatical or not. In the subject selection task, participants were required to choose the subject out of the two NPs in a test sentence; a third choice 'cannot decide' was also allowed. To ensure that the participants were paying attention to the subject selection task, for half of the tasks, the presented order of the two NPs in the question was the reverse of their positions in the test sentences. Before completing the survey, participants were instructed that there were no standard answers to the tasks and that they should base their answers on their first impressions.

2.4 Data analysis

The experimental design of the present study led to data with multiple measures per participant per condition. To deal with this kind of data, mixed-effects models with crossed random effects for participants and items (i.e., test sentences) were utilised in the analysis (Baayen et al. 2008).

For the data from the acceptability judgement task, a mixed-effects linear regression model was applied in the analysis (model 1). For the data from the grammaticality judgement task, a mixed-effects logistic regression model was applied in the analysis (model 2). For the data from the subject selection task, two separate mixed-effects logistic regression models were applied in the analysis: one for the analysis of whether decisions on the subject were made (model 3) and the other one for the analysis of which NP was chosen, given that a decision on the subject was made (model 4).

The structure of fixed effects was the same across all models. This included the within-participants factors *animacy contrast* and *word order*, which cross-classify the test sentences, the between-participants factor *group*, which indicates the group of a participant (younger or older), plus all the interactions between the three factors. The deviation coding scheme was chosen for all the factors. With regard to random effects, we employed the maximal random effects structure as proposed by Barr et al. (2013). For the by-participant random effects, we tried to include the random intercept, the random slopes for the factor *animacy contrast* and *word order* as well as the random slope for their interaction. For the by-item random effects, we tried to include the random

intercept and the random slope for the factor *group*. Model 2 successfully converged with the maximal specification of random effects, while other models required different degrees of simplification. Following Barr et al. (2013), we prioritised the preservation of random slopes during model simplification. The by-item random intercept in model 1 was dropped. The by-item random correlation of the intercept and the slope was removed in model 3 and model 4. In model 4, the by-participant random slope for the interaction between *animacy contrast* and *word order* was also dropped due to the insufficient data.

3 Results

3.1 Acceptability judgement

An overall analysis of variance on model 1 revealed a three-way interaction of *animacy contrast* × *word order* × *group* ($F(2, 66.7) = 6.77$, $p = .002$). Therefore, the results could not be simply described by any constant main effect. To gain a detailed picture of the results, we estimated the mean acceptability with 95% confidence intervals (CIs) for each condition (see Fig. 2 and Table 3). Degrees of freedom of the estimates were determined by the Satterthwaite's approximation (Kuznetsova et al. 2016). For pairwise comparison of more than two levels, *p* values (*p.adj*) and CIs were adjusted with Tukey's method.

As can be seen in Fig. 2, although the younger and older groups shared a similar pattern of the estimates for the [−animacy contrast] sentences (the right panel of Fig. 2), their patterns of estimates seemed to diverge for the [+animacy contrast] sentences (the left panel of Fig. 2).

We will first examine the estimates under the [+animacy contrast] condition. As expected, in both groups, the acceptability of the canonical SVO sentences was very close

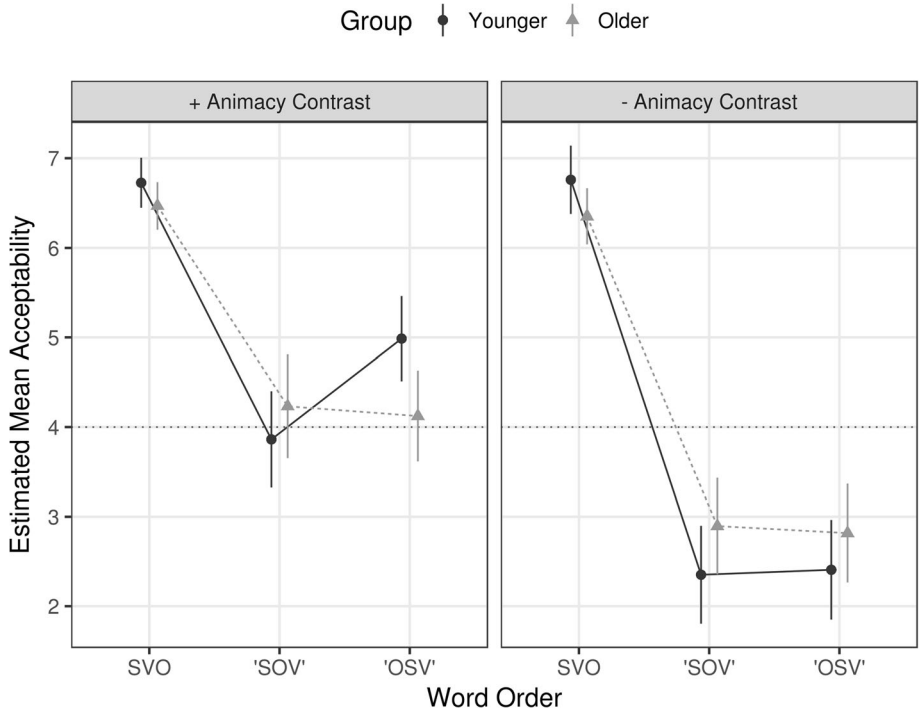

Fig. 2 Estimated acceptability

Table 3 Estimated acceptability

Animacy contrast	Word order	Younger		Older	
		Estimate	95% CI	Estimate	95% CI
+	SVO	6.73	[6.45, 7.01]	6.47	[6.20, 6.73]
+	'SOV'	3.86	[3.32, 4.40]	4.23	[3.65, 4.81]
+	'OSV'	4.99	[4.51, 5.46]	4.12	[3.62, 4.63]
–	SVO	6.76	[6.38, 7.14]	6.35	[6.04, 6.66]
–	'SOV'	2.35	[1.80, 2.90]	2.90	[2.36, 3.43]
–	'OSV'	2.41	[1.85, 2.96]	2.82	[2.27, 3.37]

to the highest score of 7 (younger 6.73, 95% CI [6.45, 7.01]; older 6.47, [6.20, 6.73]). The estimated difference between the two groups was 0.26, [– 0.09, 0.61], which did not reach statistical significance, $t(61.8) = 1.48$, $p = .144$. The estimated acceptability of the 'SOV' sentences was also very similar in the two groups. The estimates were both near to the neutral score of 4 (younger 3.86, [3.32, 4.40]; older 4.23, [3.65, 4.81]), and the estimated between-group difference was not statistically significant (– 0.37, [– 1.15, 0.41], $t(53.9) = 0.95$, $p = .346$). It was in the 'OSV' sentences where the two groups diverged noticeably. In the older group, the acceptability of the 'OSV' sentences was close to the neutral score of 4 (4.12, [3.62, 4.63]). Also, the estimate did not deviate significantly from that of the 'SOV' sentences in the same group (estimated 'SOV' – 'OSV' difference in the older group: 0.11, [– 0.51, 0.73], $t(59.9) = 0.42$, $p.\ adj = .908$). In the younger group, by contrast, the acceptability of the 'OSV' sentences tended reasonably toward acceptance, with an estimated mean of 4.99, [4.51, 5.46]. The acceptability of the 'OSV' sentences was also significantly higher than the 'SOV' sentences in the younger group, by an estimated difference of 1.12, [0.51, 1.73], $t(84.7) = 4.41$, $p.\ adj < .001$. The difference of the acceptability of the 'OSV' sentences between the two groups was further confirmed by a between-group comparison. The acceptability in the younger group was higher than that in the older group by a statistically significant difference of 0.86, [0.18, 1.54], $t(54.5) = 2.54$, $p = .014$.

Let us continue to examine the estimates under the [–animacy contrast] condition. The patterns of the estimated acceptability were similar between the two groups. In general, the acceptability of the canonical SVO sentences was near the highest score of 7, and the acceptability of the non-canonical 'SOV' and 'OSV' sentences was below the neutral score of 4. Recall that the nouns in the [–animacy contrast] sentences were both human names, and consequently, the derived non-canonical 'SOV' and 'OSV' sentences were both animate-animate-verb (AAV) strings. Thus, they were merely two variants of the same kind, i.e., the AAV strings. Therefore, in principle, no difference should be found in the acceptability of the 'SOV' and 'OSV' sentences. This was confirmed by the analysis. In the younger group, the estimated difference between the acceptability of the 'SOV' and 'OSV' sentences was – 0.06, 95% CI [– 0.63, 0.52], $t(126.2) = 0.23$, $p.\ adj = .971$. In the older group, the estimated difference was 0.08, [– 0.32, 0.48], $t(99.7) = 0.47$, $p.\ adj = .887$. Neither of these differences reached statistical significance. Furthermore, there was no significant between-group difference found in the 'SOV' or 'OSV' sentences. The estimated younger – older group difference at 'SOV' was – 0.54, [– 1.26, 0.17], $t(58.9) = 1.52$, $p = .134$. The estimated group difference at 'OSV' was – 0.41, [– 1.14, 0.32], $t(58.6) = 1.12$, $p = .267$. Neither was statistically significant.

Under the [−animacy contrast] condition, however, there was one seemingly problematic result coming out from the SVO sentences. Although in both groups, the SVO sentences were estimated to have very high acceptability (younger 6.76, 95% CI [6.38, 7.14]; older 6.35, [6.04, 6.66]), the estimated younger − older group difference was found to be statistically significant, 0.41, [0.01, 0.80], $t(77.1) = 2.04$, $p = .045$. Does this imply that the two groups showed a substantial difference in their acceptability of the canonical SVO sentences? We doubt it for four reasons. First, consider that the point estimate of the between-group difference (0.41) was below 0.5, which was relatively small in the context of a 7-point Likert scale, and that the 95% CI only marginally excluded zero ([0.01, 0.80]). We think it would be too soon to jump to the conclusion that the difference was of practical significance. Second, the estimated difference suggested that, for the SVO sentences under the [−animacy contrast] condition, the acceptabilty in the older group was slightly but significantly lower than the younger group. On the other hand, as described above, under the [+animacy contrast] condition, the estimated younger − older difference for the SVO sentences was of no statistical significance, 0.26, [− 0.09, 0.61], $t(61.8) = 1.48$, $p = .144$. Although they differed in statistical significance, when the two between-group differences in SVO sentences (one under the [+animacy contrast] condition, the other under the [−animacy contrast] conditions) were contrasted with each other, no significant difference was confirmed, 0.15, [− 0.14, 0.43], $t(116.3) = 1.03$, $p = .304$. This means that the slight between-group difference in SVO sentences under the [−animacy contrast] condition is hardly greater than that under the [+animacy contrast] condition. Third, as previously shown in Table 2, the amount of test sentences for the [−animacy contrast] condition (i.e., 30) was comparatively less than the [+animacy contrast] ones (i.e., 90). This would contribute relatively more sampling error to the results under the [−animacy contrast] condition, which might account for the unexpected statistical significance here. Fourth and finally, in the grammaticality judgement task, as detailed in the next section, there was no significant between-group difference found in the SVO sentences under both animacy conditions. Judging from the reasons above, we think this minor between-group difference in the SVO sentences under the [−animacy contrast] condition was more likely due to sampling error than a real linguistic difference in the population.

In summary, regardless of the conditions of animacy, the acceptability of the SVO sentences was close to the highest score in both groups, which was not surprising because SVO is the canonical word order in Chinese. Under the [−animacy contrast] condition, the non-canonical 'SOV' and 'OSV' sentences (i.e., the AAV strings) were generally not acceptable in both groups. Under the [+animacy contrast] condition, the acceptability of the 'SOV' sentences (i.e., the AIV strings) was neutral in both groups. However, the acceptability of the 'OSV' sentences (i.e., the IAV strings) was significantly different in the two groups. In the older group, the acceptability of the 'OSV' sentences was as neutral as the 'SOV' sentences. In the younger group, by contrast, the 'OSV' sentences were biased to be acceptable, and the acceptability was significantly higher than that of the 'SOV' sentences in the same group. Furthermore, a between-group difference for the 'OSV' sentences under the [+animacy contrast] condition was confirmed, with the acceptability in the younger group being significantly higher than that in the older group. It can be concluded that the older native speakers showed lower

acceptance toward the 'OSV' sentences under the [+animacy contrast] condition (i.e., the IAV strings) than the younger speakers.

3.2 Grammaticality judgement

An overall analysis of deviance on model 2 revealed a three-way interaction of *animacy contrast* × *word order* × *group* ($\chi^2(2) = 17.79$, $p < .001$). Therefore, the results could not be simply described by any constant main effect. To gain a detailed picture of the results, we estimated the mean probabilities of sentences being judged as grammatical (the 'grammatical probability') with 95% CIs for each condition (see Fig. 3 and Table 4). Because of the nature of logistic regression, the estimates were originally in logarithmic scale and they represented the log odds of probabilities. To make them intuitive to understand, we transformed the log odds to probabilities with the inverse logit function. For the estimates on differences, however, we left them in their original logarithmic scale, as they were only used for statistical significance tests here. For pairwise comparison of more than two levels, *p* values (*p.adj*) and CIs were adjusted with Tukey's method.

First, let us examine the results for the SVO sentences. As shown in Fig. 3 and Table 4, the point-estimated probabilities of the SVO sentences being judged grammatical were above 99% for both groups under both animacy conditions, with a slight variance in their 95% CIs. There was no significant between-group difference for the SVO sentences in either animacy condition (estimated younger – older group difference under the [+animacy contrast] condition 0.21, 95% CI [– 1.05, 1.47], $z = 0.33$, $p = .742$; under the [–animacy contrast] condition 1.46, [– 0.52, 3.45], $z = 1.45$, $p = .147$). These

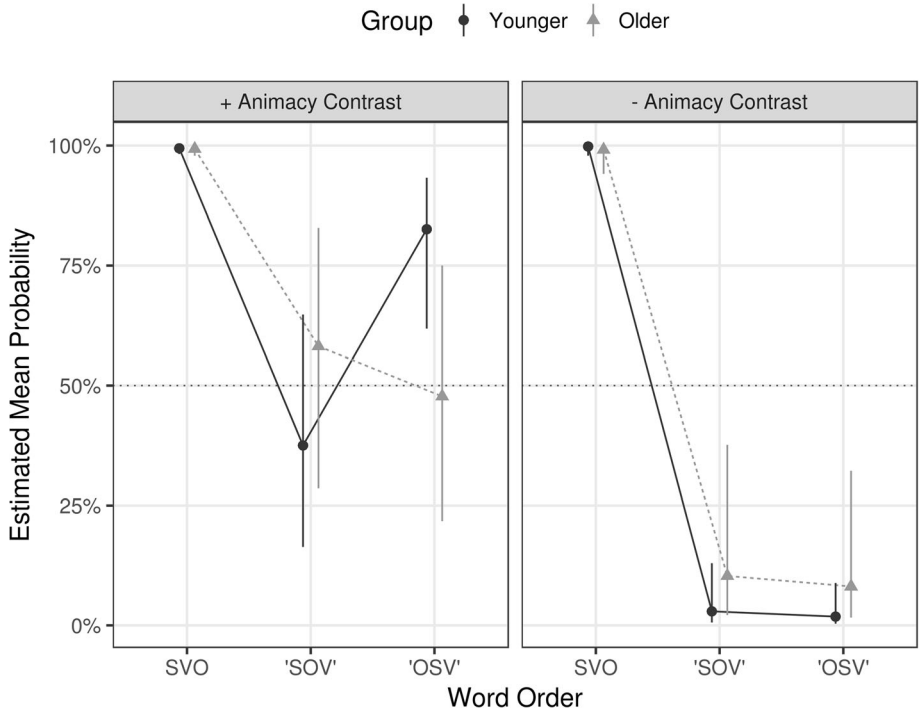

Fig. 3 Estimated grammatical probability

Table 4 Estimated grammatical probability

Animacy contrast	Word order	Younger		Older	
		Estimate	95% CI	Estimate	95% CI
+	SVO	99.4%	[98.2, 99.8]	99.3%	[97.9, 99.8]
+	'SOV'	37.5%	[16.4, 64.8]	58.2%	[28.6, 82.9]
+	'OSV'	82.6%	[61.8, 93.3]	47.7%	[21.8, 75.0]
–	SVO	99.8%	[97.9, 100.0]	99.1%	[94.1, 99.9]
–	'SOV'	2.9%	[0.6, 13.0]	10.3%	[2.1, 37.6]
–	'OSV'	1.8%	[0.4, 8.8]	8.1%	[1.6, 32.2]

results are expected since SVO is the canonical word order. In the following, we discuss the results for the non-canonical sentences.

Under the [–animacy contrast] condition, the non-canonical 'SOV' and 'OSV' sentences (i.e., the AAV strings) were generally judged as ungrammatical, as their estimated grammatical probabilities were considerably low. The point-estimated probabilities were all below 11%. The upper limits of the 95% CIs were all below the 50% point; therefore, the tendency of ungrammatical judgement was not merely by chance. Although the estimated probabilities in the older group were relatively higher than those in the younger group, the margins of the 95% CIs in the older group were also appreciably wider, indicating larger uncertainty in the estimates. No significant difference was found in the between-group comparisons for either the 'SOV' or the 'OSV' sentences (estimated younger – older group difference at 'SOV' – 1.34, 95% CI [– 3.41, 0.73], $z = 1.27$, $p = .205$; at 'OSV' – 1.56, [– 3.64, 0.53], $z = 1.47$, $p = .143$). Also, there was no significant difference between the two variants of AAV strings in both groups (estimated 'SOV' – 'OSV' difference in the younger group 0.48, [– 1.07, 2.03], $z = 0.73$, p. adj = .748; in the older group 0.26, [– 1.04, 1.57], $z = 0.47$, p. adj = .884).

Under the [+animacy contrast] condition, the grammatical probability of the 'SOV' sentences (i.e., the AIV strings) was estimated to be near 50% in both the younger and the older groups. The 95% CIs were broad, indicating a high level of uncertainty in the estimates, which in turn indicated a high individual variance in the judgement. Neither of the CIs excluded the 50% point, therefore these probabilities could not be considered to be beyond the chance level. Although the point-estimated probabilities differed to some extent in the two groups, no significant between-group difference was found (estimated younger – older group difference – 0.84, 95% CI [– 2.48, 0.80], $z = 1.01$, $p = .314$). These results are well understandable: with the acceptability being neutral, but with the sentences needing to be dichotomously judged as grammatical or ungrammatical, the overall grammatical probability of the 'SOV' sentences will average around 50%, with high variance among the judgements.

For the 'OSV' sentences under the [+animacy contrast] condition (i.e., the IAV strings), again, a remarkable divergence was observed between the two groups. In the older group, the point-estimated grammatical probability was very close to 50% (47.7%). The 95% CI was broad, and did not exclude the 50% point ([21.8, 75.0]). Thus, the probability could not be considered to be beyond the chance level. Also, there was no significant difference between the 'OSV' and 'SOV' sentences in the older group (estimated 'SOV' – 'OSV' difference 0.42, 95% CI [– 0.82, 1.65], $z = 0.80$, p. adj = .705). In the younger group, by contrast, the grammatical probability was fairly high, with a

point estimate of 82.6%. The 95% CI was relatively narrower, and excluded the 50% point by a reasonable margin ([61.8, 93.3]). This means that the tendency of grammatical judgement was significantly above chance. In the younger group, the grammatical probability for the 'OSV' sentences was also significantly higher than the 'SOV' sentences (estimated 'OSV' – 'SOV' difference 2.07, [0.93, 3.20], $z = 4.26$, $p.$ adj < .001). Moreover, the estimated younger – older group difference was shown to be significant (1.65, [0.08, 3.21], $z = 2.06$, $p = .039$). As demonstrated by the results, the younger group judged the 'OSV' sentences under the [+animacy contrast] condition (i.e., the IAV strings) as grammatical at a considerably high probability, whereas in the older group, the grammatical probability of these sentences was merely at the chance level.

In general, the results of the grammaticality judgement task were in line with the results of the acceptability judgement task.

3.3 Subject selection

As mentioned previously, the analysis for the data of the subject selection task was done with two separate mixed-effects logistic regression models, i.e., model 3 and model 4. Model 3 was intended to analyse whether decisions on the subject were made. Specifically, it was used to analyse the 'decision probability', which represents the probability of not choosing the 'cannot decide' option. Model 4 was intended to analyse which NP was chosen, given that a decision on the subject was made. Specifically, it was used to analyse the 'NP1 probability', which represents the conditional probability of the first NP being chosen as the subject. In contrast to models 1–3, which worked with the full set of the final data, model 4 only worked with a subset of the data, i.e., the responses where decisions on the subject were made.

We will first examine the results of model 3.

3.3.1 Decision probability

An overall analysis of deviance was performed on model 3. Although the three-way interaction *animacy contrast* × *word order* × *group* did not reach significance ($\chi^2(2) = 1.44$, $p = .487$), two significant two-way interactions were identified: the *animacy* × *group* interaction ($\chi^2(1) = 6.43$, $p = .011$) and the *animacy* × *word order* interaction ($\chi^2(2) = 18.95$, $p < .001$). Because of the interactions, the results could not be simply described by any constant main effect. To gain a detailed picture of the results, we estimated the mean decision probabilities with 95% CIs for each condition (see Fig. 4 and Table 5). As with model 2, we had transformed these estimates from log odds to probabilities for intuitive interpreting, but left the estimates on differences in logarithmic scale for simplicity. For pairwise comparison of more than two levels, p values (*p.adj*) and CIs were adjusted with Tukey's method.

Under the [+animacy contrast] condition, the decision probabilities in both groups were close to the ceiling for all the levels of *word order*, with most of the point estimates being above 99%. In the younger group, while there was no significant difference between the estimated decision probabilities for the canonical SVO and the non-canonical 'OSV' sentences (1.31, 95% CI [– 1.20, 3.82], $z = 1.22$, $p.$ *adj* = .440), the estimate for the 'SOV' sentences (95.5%, [83.1, 98.9]) was slightly lower than the other two cases (estimated SVO – 'SOV' difference 3.46, [1.17, 5.75], $z = 3.54$, $p.$ *adj* = .001;

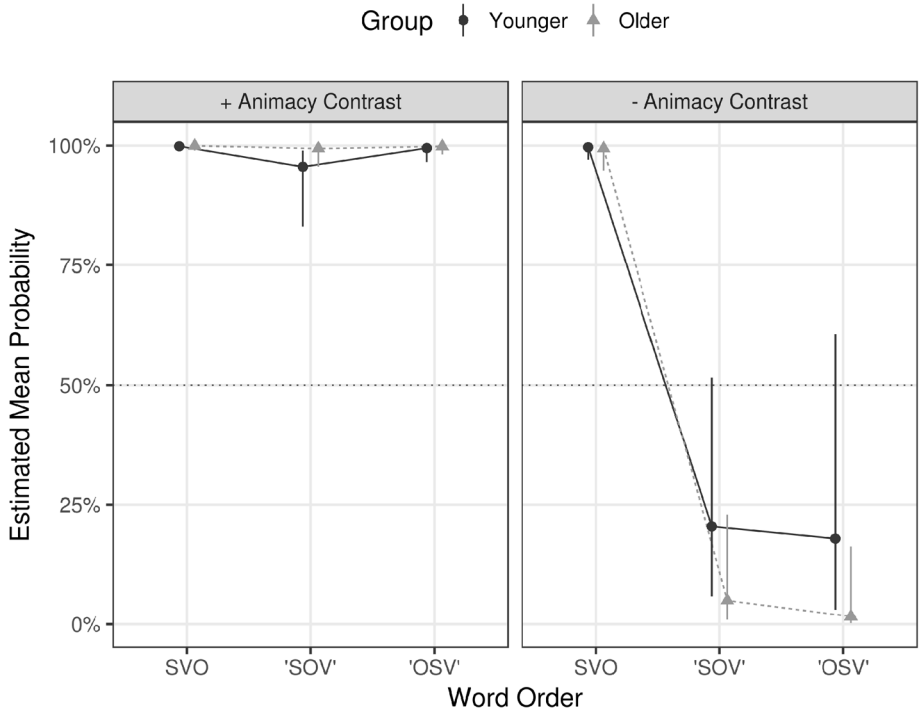

Fig. 4 Estimated decision probability

estimated 'OSV' – 'SOV' difference 2.15, [0.14, 4.17], $z = 2.50$, $p.\ adj = .033$). For non-canonical sentences, if decisions on the subject were made solely based on the animacy contrast, then there should be no significant difference between the decision probabilities of the 'SOV' and 'OSV' sentences, since they only swapped the positions of the two nouns. The moderately lower decision probability for the 'SOV' sentences in the younger group might be an indication that the group has a preference for the OSV word order. As a consequence, a small fraction of the responses might have chosen the 'cannot decide' option because these responses had considered SOV as just uninterpretable. In the older group, on the other hand, there was no significant difference among the decision probabilities for the SOV, 'SOV', and 'OSV' sentences (estimated SVO – 'SOV' difference 2.42, [– 1.40, 6.25], $z = 1.49$, $p.\ adj = .297$; estimated SVO – 'OSV' difference 1.16, [– 2.72, 5.05], $z = 0.70$, $p.\ adj = .763$; and estimated 'SOV' – 'OSV' difference – 1.26, [– 4.03, 1.51], $z = -1.07$, $p.\ adj = .533$).

Table 5 Estimated decision probability

Animacy contrast	Word order	Younger		Older	
		Estimate	95% CI	Estimate	95% CI
+	SVO	99.9%	[99.1, 100]	99.9%	[98.9, 100]
+	'SOV'	95.5%	[83.1, 98.9]	99.4%	[95.5, 99.9]
+	'OSV'	99.5%	[96.5, 99.9]	99.8%	[98.2, 100]
–	SVO	99.7%	[97.1, 100]	99.4%	[94.8, 99.9]
–	'SOV'	20.5%	[5.9, 51.6]	5.0%	[0.9, 23.0]
–	'OSV'	17.9%	[3.0, 60.6]	1.6%	[0.1, 16.3]

Under the [–animacy contrast] condition, the decision probabilities of the SVO sentences were also extremely high, with point estimates above 99%. However, there was a huge drop in the estimates for the non-canonical 'SOV' and 'OSV' sentences (i.e., the AAV strings) in both groups. The point estimates in the younger group were 20.5% and 17.9%, with the upper limits of 95% CIs being no higher than 61%. The point estimates in the older group were 5.0% and 1.6%, with the 95% CIs being far below 50%. This demonstrates a difficulty in choosing subjects in AAV strings. Although the estimates in the younger group appeared to be higher than the older group, the 95% CIs were also much broader, indicating larger uncertainty in the estimates. No significant between-group difference was found (estimated younger – older group difference at 'SOV' 1.59, 95% CI [– 0.57, 3.76], $z = 1.45$, $p = .148$; at 'OSV' 2.60, [– 0.44, 5.63], $z = 1.68$, $p = .093$). There was also no significant difference between the two variants of AAV strings in both groups (estimated 'SOV' – 'OSV' difference in the younger group 0.16, [– 0.92, 1.25], $z = 0.35$, $p = .933$; in the older group 1.17, [– 0.49, 2.83], $z = 1.65$, $p = .225$).

The high decision probabilities for the 'SOV' and 'OSV' sentences under the [+animacy contrast] condition (i.e., AIV and IAV strings) show that people have little to no difficulty in deciding the subject in non-canonical sentences with animacy contrast. In sharp contrast with this, the decision probabilities for the non-canonical sentences under the [–animacy contrast] condition (i.e., AAV strings) were considerably low, indicating a difficulty in analysing the syntactic and thematic roles in such strings.

3.3.2 NP1 probability

An overall analysis of deviance was performed on model 4. All the interactions that contained the factor *group*, along with the main effect of the *group* itself, were shown to be non-significant (*animacy contrast* × *word order* × *group* $\chi^2(2) = 2.83$, $p = .243$; *animacy contrast* × *group* $\chi^2(1) = 3.11$, $p = .078$; *word order* × *group:* $\chi^2(2) = 1.16$, $p = .560$; *group* $\chi^2(1) = 0.30$, $p = .581$). Thus, it can be inferred that there was no significant between-group difference in any condition. There was, however, a significant two-way interaction of *animacy contrast* × *word order* ($\chi^2(2) = 60.64$, $p < .001$). Because of the interaction, the results still could not be simply described by any constant effect of *animacy contrast* or *word order.* To gain a detailed picture of the results, we estimated the mean NP1 probabilities with 95% CIs for each condition (see Fig. 5 and Table 6). As with model 2 and model 3, we had transformed these estimates from log odds to probabilities for intuitive interpreting but left the estimates on differences in logarithmic scale for simplicity. For pairwise comparison of more than two levels, *p* values (*p.adj*) and CIs were adjusted with Tukey's method.

Under the [+animacy contrast] condition, the NP1 probabilities of the SVO and 'SOV' sentences were close to the ceiling in both groups, with all their point estimates being above 97%. No statistically significant difference was found between the SVO and 'SOV' sentences in both groups (estimated SOV – 'OSV' difference in the younger group 1.28, [– 0.42, 2.99], $z = 1.76$, *p*. adj = .182; in the older group 0.03, [– 2.22, 2.29], $z = 0.03$, *p*. adj = .999). The high NP1 probability was not surprising for the non-canonical 'SOV' sentences (i.e., the AIV strings), because the animate nouns, which were also the original subjects, were positioned at the beginning of these sentences. On the other hand, the NP1 probability of the 'OSV' (i.e., the IAV

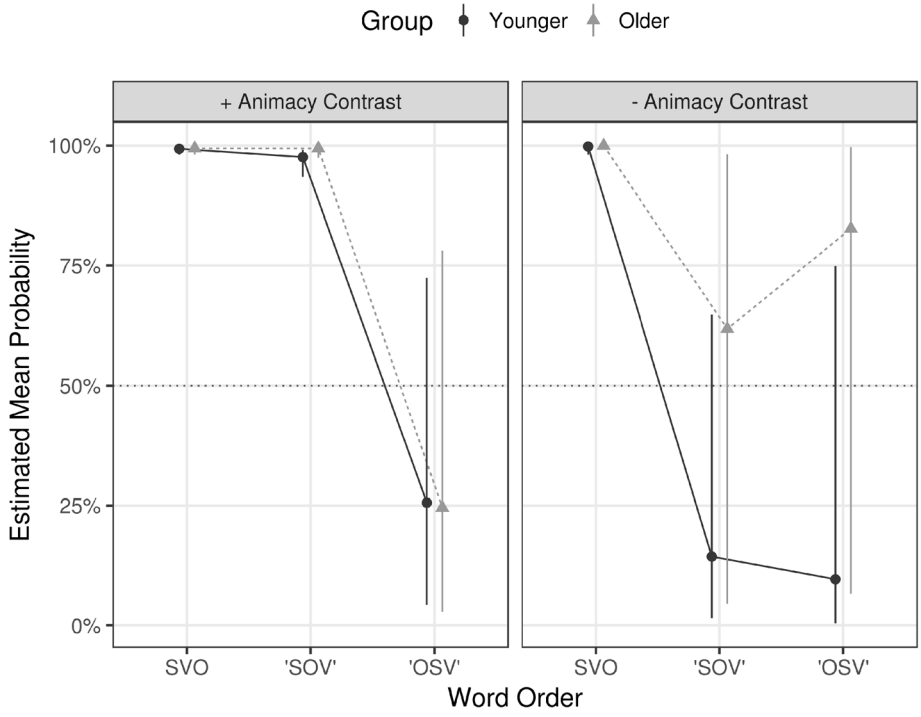

Fig. 5 Estimated NP1 probability

strings) sentences was remarkably low. In the younger group, the estimate was 25.6%, 95% CI [4.3, 72.4]; in the older group, the estimate was 24.6%, [2.9, 78.2]. Because the animate nouns (also the original subjects) were positioned as the second nouns in the 'OSV' sentences, the low point estimates of the NP1 probabilities were expected. However, it is interesting that the point estimates did not hit the bottom, and the 95% CIs were broad. This implies that there was a relatively high individual variance in the sample, and a reasonable amount of responses had taken the first, inanimate nouns as the 'subject'. The almost identical estimates in the younger and older groups means that this was the same situation for both groups. The explanation for this outcome might largely lie in the very term 'subject' that we used in this task. The distinction of subject and topic is rather controversial in Chinese (Shi 2001), and some renowned linguists such as Zhu Dexi even consider all sentence-initial nouns as the subject (Zhu 朱德熙 1982: 95–96). It is possible

Table 6 Estimated NP1 probability

Animacy contrast	Word order	Younger		Older	
		Estimate	95% CI	Estimate	95% CI
+	SVO	99.3%	[98.1, 99.8]	99.4%	[98.1, 99.8]
+	'SOV'	97.6%	[93.5, 99.1]	99.4%	[97.4, 99.9]
+	'OSV'	25.6%	[4.3, 72.4]	24.6%	[2.9, 78.2]
–	SVO	99.8%	[98.1, 100]	100%	[99.6, 100]
–	'SOV'	14.4%	[1.5, 64.8]	61.8%	[4.6, 98.2]
–	'OSV'	9.7%	[0.4, 74.9]	82.7%	[6.7, 99.7]

that some of the responses had confused the concept of subject with the concept of topic and, therefore, took the topicalized object as the 'subject'.

Under the [–animacy contrast] condition, the NP1 probability of the SVO sentences was extremely high, with the point estimates being above 99%. This was expected as SVO is the canonical word order. For the non-canonical 'SOV' and 'OSV' sentences (i.e., the AAV strings), there were great uncertainties in the estimates. In the younger group, the point estimates were below 15%, but the upper limits of the 95% CIs could be as high as 74.9%. In the older group, the point estimates were 61.6% and 82.7%, but the 95% CIs ranged from below 10% to above 98%. In an AAV string, the two nouns were both human names; therefore in principle, they were equally likely to be chosen as the subject. The point-estimated NP1 probabilities suggested that based on the sample, there might be an NP2 preference in the younger group and an NP1 preference in the older group. However, the 95% CIs were so broad in the two groups that no definite conclusions could be drawn. Indeed, there was no statistically significant between-group difference in any case, as already shown by the overall analysis of deviance. Also, in both groups, no significant difference between the two variants of the AAV strings was found (estimated 'SOV' – 'OSV' difference in the younger group 0.45, [– 2.87, 3.78], $z = 0.32$, $p.$ adj = .946; in the older group – 1.09, [– 5.76, 3.58], $z = 0.55$, $p.$ adj = .849). The high level of uncertainty in the estimates of the 'SOV' and 'OSV' sentences could be an indication of a great degree of individual variance in the interpretation of AAV strings. Meanwhile, the insufficiency of data was also a cause. As mentioned previously, the amount of available data for model 4 was determined by whether a decision on the subject was made. Recall that the AAV strings were the only case where the decision probability was remarkably low. This means that for these sentences, only a small portion of the data were available for the analysis of NP1 probability. Take the 'SOV' sentences for example. In the raw data, only 34% (99 out of 290) of the responses was available from the younger group, and only 22% (51 out of 230) was available from the older group. The small size of the data would lead to a larger margin of error in the estimates of the AAV strings.

4 Discussion

This preliminary study examined the acceptability of non-canonical word orders in Chinese with a group of younger adults and a group of older adults. We found that the two groups showed congruent results under almost all our experimental conditions, except for the IAV strings.

The baseline SVO sentences were, with no surprise, judged as highly acceptable in both groups. Animacy contrast did not manifest any influence on the results of SVO sentences. Irrespective of the condition of animacy contrast, SVO sentences were estimated to be between 6 and 7 points in our 7-point scale acceptability judgement task. In our grammaticality judgement task, SVO sentences were estimated to be judged as grammatical over 99% of the time. The results from our subject selection task further guaranteed that people have no difficulty in analysing the syntactic relations in SVO sentences. The decision probability (i.e., the probability of not choosing the 'cannot decide' option) was estimated to be above 99% in all the cases of SVO, and over 99% of the time the first noun would be chosen as the subject. All the estimates have come

with remarkably narrow margins of error when compared to non-canonical test sentences, which is evidence for a high consistency among individuals.

Prior work with theoretical and experimental approaches has demonstrated that the SOV word order is subject to more restrictions and is more difficult to process compared to the OSV word order. For AAV strings (the 'SOV' and 'OSV' strings under the [−animacy contrast] condition), where no animacy contrast exists, theoretical studies such as Hou (1979), Lu (1994), and Qu (1994) have made strong claims that OSV is the sole interpretation. Nevertheless, the experimental studies of Miao (1981), Li et al. (1992), and Li et al. (1993) have shown that the OSV interpretation of AAV strings is far less dominant than expected. More interestingly, their results in fact contradicted each other, with Miao (1981) suggesting a SOV preference and Li et al. (1992) and Li et al. (1993) suggesting an OSV preference. We postulated that, contrary to previous studies, there might not exist a default interpretation for non-canonical sentences when the animacy of the two nouns are perfectly uniform, such as the AAV strings in the present study. Our results have supported this hypothesis. AAV strings were the only case that the decision probability was estimated to be below 21%, while other cases all had the decision probability above 95%. Even for the responses where decisions on the subject were made, a great level of inconsistency existed in both groups, rendering it impossible to infer any structural preference. In addition, AAV strings are generally regarded as unacceptable. In both groups, the acceptability of AAV strings was estimated to be between 2 and 3, which is significantly below the neutral score of 4. And it is well beyond the chance level that AAV strings would be judged as ungrammatical. The possibility of AAV strings being judged as grammatical was estimated to be below 11%. Because the decision probability was considerably low for AAV strings, the poor acceptability of AAV strings should be a direct result of their uninterpretable status.

In regard to non-canonical sentences with animacy contrast, recall that the 'SOV' and 'OSV' strings under the [+animacy contrast] condition in the present study are equivalent to AIV and IAV strings respectively. For these cases, the estimated decision probabilities of the subject selection task were all above 95%, which indicates that the syntactic and thematic relations are resolvable in AIV and IAV strings. However, being interpretable is not equal to being fully acceptable. In our data, the acceptability of AIV and IAV strings was always estimated to be lower than that of canonical SVO sentences. For AIV strings, the acceptability was only around the neutral point in both groups (3.86, 95% CI [3.32, 4.40] in the younger group; 4.23, [3.65, 4.81] in the older group). The probability of AIV strings being judged as grammatical was also not significantly different from the chance level of 50% in both groups (37.5%, [16.4, 64.8] in the younger group, 58.2% [28.6, 82.9] in the older group). The younger group and the older group did not significantly differ from each other either by the measurement of acceptability or by the measurement of grammaticality.

The only case where the younger group and the older group differed significantly is in the acceptability and grammaticality of IAV strings. Crucially, we found that the older group expressed significantly lower acceptance toward IAV strings when compared to the younger group. The acceptability of IAV strings in the younger group was biased to be acceptable (4.99, 95% CI [4.51, 5.46]), and it was significantly higher than the acceptability of AIV strings in the same group (p.adj< .001). In the older group, however, the acceptability of IAV strings had dropped to the neutral level (4.12, [3.62, 4.63]), and it did not

significantly differ from the acceptability of AIV strings (p.adj = .908). The different tendencies between the two groups were also confirmed by the results of the grammaticality judgement task. In the younger group, the probability that IAV strings are judged as grammatical was as high as 82.6%, [61.8, 93.3], which was also significantly higher than the grammatical probability of AIV strings in the same group (p.adj < .001). By contrast, the probability of IAV strings being judged as grammatical was only around the chance level in the older group (47.7%, [21.8, 75.0]), and it did not significantly deviate from that of AIV strings (p.adj = .705). Moreover, between-group comparison reaffirmed that the acceptability of IAV strings in the younger group was significantly higher than that of the older group (p = .014). And it was also true for the between-group comparison of grammaticality (p = .039). Although the ambiguity of the concept of subject and topic in Chinese (Shi 2001) might have caused some confusion to the participants in both groups when performing the subject selection task, previous studies have clearly shown that AIV strings are basically interpreted as SOV, and IAV strings are basically interpreted as OSV (Miao 1981; Li et al. 1992; Li et al. 1993). In addition, the AIV and IAV strings in the present study were directly derived from corresponding SVO sentences in which the animacy contrast exists and the thematic roles of nouns are not exchangeable. Therefore, the acceptability of the AIV and IAV strings could be interpreted as the acceptability of SOV and OSV.

Previous studies (Li et al. 1992; Li et al. 1993; Lu 1994; Ernst and Wang 1995; Paul 2005; Huang et al. 2009; Wang et al. 2012; Wang 王路明 2015) have illustrated that there is a preference for OSV over SOV in Chinese. However, few studies have considered that the preference for a particular word order might systematically vary with certain individual factors. We found that while a bold advantage of OSV over SOV existed in the younger group, it had ceased to be visible in the older group. Although uncontrolled factors such as the difference of educational backgrounds between the two groups[3] might play a role, we think it is less likely because the stimuli in this study are very simple, and non-canonical sentences with bare objects are more frequent in spoken Chinese (Sun and Givón 1985). Another concern might be the fact that the participants in the younger group were also learners of Japanese with JLPT N1 level (the highly proficient level). However, no evidence for L2 influence was found in the present study because SOV is far more frequent than OSV in Japanese, with a reported SOV:OSV ratio of 17:1 (Kuno 1978). If the younger group's perception of word orders in Chinese were influenced by Japanese, higher acceptability for the AIV (SOV) strings should be observed, which as shown by our data is not true. Judging from the above, we consider this difference to be age-related.

Now the question comes that why age-related differences were found only in the OSV-interpreted IAV strings but not the AIV and AAV strings. First, we think that there might be little or no room for age-related changes in the acceptability of AIV and AAV strings (i.e., floor effects). The results of the present study strongly indicate that the acceptance of non-canonical word orders in Chinese is dependent on the resolvable syntactic relations of the pre-verbal nouns, which in turn rely on the existence of semantic cues such as animacy contrast. For AAV strings, the lack of animacy contrast hinders a comprehender from further interpreting these strings. This directly leads to the poor acceptability of AAV strings and leaves no room for age to play a role. For AIV strings, the animacy contrast makes it possible to analyse the syntactic relations of

the pre-verbal nouns. Thus, they should at least have a random chance to be judged as acceptable. Meanwhile, the use of the SOV word order is known to be rather restricted in Chinese. As we have seen, the acceptability and grammaticality of AIV strings in the younger group did not go beyond the neutral or chance level. It is possible that the overall neutral judgement is both the least and the most that AIV strings can have; as a result, age can impose little effects on the acceptability of these strings. The IAV strings are similar to the AIV strings in the sense that the syntactic relations of the pre-verbal nouns can be resolved with the help of animacy cues. An important difference, however, is that the OSV word order is less restricted than the SOV word order both syntactically and semantically in Chinese. In addition, OSV has also been demonstrated to be easier to process than SOV. Because of the properties of the OSV word order, the OSV-interpreted IAV strings should have the potential to be judged as more acceptable. Therefore, the possible overall acceptability of IAV strings is not constrained to the neutral level but can range from merely neutral to somewhere on the more acceptable side.

The striking difference in the acceptability of IAV strings between the younger adults and the older adults suggests that there must be some changes in the use or perception of OSV sentences between the two age groups. One possible change is that younger adults might have more frequent use of the OSV word order than older adults, thus giving rise to the higher acceptability in the group. As discussed by Wagner (2012), younger adults' language is relatively free from normative pressures while older adults might employ more standard forms of speech. The OSV word order is an optional construction in spoken Chinese which involves the topicalization of the object. The subjective acceptability of OSV can depend on how frequently speakers from an age group use topicalized OSV sentences in their linguistic environment. Do younger adults use topicalized sentences more frequently, or do older adults adhere more to the canonical word order? For now, the lack of corpus studies on the language use in younger and older adults in the Chinese linguistic literature means that we have no direct answer. However, we should not refrain from consulting studies of other languages. In spoken French, for example, the left and right dislocations of a noun or pronoun serve as devices to introduce topics (Ashby 1982; Ashby 1988; Katz and Blyth 2007). Ashby's (1982) corpus-based study found that both the left and right dislocations are significantly more frequent in the discourse of younger speakers (60%) than that of older speakers (32%). This example of spoken French shows that quantitatively different uses of the topic construction between age groups is possible. The question whether a similar situation exists in the use of the OSV word order in Chinese calls for further corpus-based investigations.

Another possibility is that the lower acceptability of OSV in the older adults might be related to cognitive ageing. Older adults are generally found to have relatively worse performance during online sentence processing (Wlotko et al. 2010), recent studies have provided compelling evidence that native Chinese speakers are no exception (Xu et al. 2017; Z. Zhu et al. 2018). Previous studies in English and Korean report that age effects are most pronounced in the processing of structures with non-canonical word orders, such as object relative clauses and object clefts in English (Stine-Morrow et al. 2000; Caplan et al. 2011) and the OSV word order in Korean (Oh et al. 2016; Sung 2017). In a similar vein, considering that OSV is a non-canonical word order in Chinese, it is possible that older native speakers of Chinese might also have more

difficulties when processing OSV sentences relative to younger adults. Although acceptability judgements can only be made after the processing of a sentence and therefore are not directly linked with processing performance, subjective acceptability and processing performance measures can often be correlated, as is evident in the data of Wang et al. (2012) and Wang 王路明 (2015). Hence, it also seems worthwhile to scrutinize whether the lower acceptability of OSV in the older adults is associated with age-related differences in the processing of OSV by using experimental methods, which has not yet been done in the literature.

5 Conclusions

The present study provides valuable data on the acceptability of non-canonical word orders among younger and older native speakers of Mandarin Chinese. Previous theoretical and experimental studies demonstrated that OSV word order is preferred over SOV word order in Chinese. However, our results suggest that this preference for OSV can be modulated by the factor of age and therefore is not invariant. Specifically, whereas OSV was judged significantly more acceptable than SOV in the younger adults, the acceptability of OSV in the older adults was only neutral and did not significantly deviated from that of SOV. While only speculations are allowed with the data available at the moment, there must be an underlying cause for this striking difference. On one hand, the difference can be an indication of a possible quantitative difference in the use of OSV between different age groups. On the other hand, it also seems possible that the reduced acceptability of OSV in the older adults can be associated with age-related differences in sentence processing. Future corpus-based and experimental work are required to examine the possible explanations proposed in the present study.

6 Endnotes

[1]We consider SVO as the only canonical word order in Chinese for its overwhelmingly high frequency in use.

[2]The prefix *xiǎo-* 'small-' in human names is used to refer to a person who is younger than the speaker with a sense of familarity (Sun 2006).

[3]It was very difficult to control the educational background of the two groups because of the difficulty of recruiting older adults with higher education experience. Due to historical reasons, many of them did not have the chance to receive higher education at their younger ages.

Abbreviations

AAV: Animate-animate-verb; AIV: Animate-inanimate-verb; CI: Confidence interval; CL : Classifier; ERP: Event-related potential; IAV: Inanimate-animate-verb; IF : Inflected future; NP: Noun phrase; OSV: Object subject verb; PERF: Perfect aspect; PF : Periphrastic future; SOV: Subject-object-verb; SVO: Subject-verb-object

Acknowledgements

We sincerely thank the two anonymous reviewers for their helpful comments.

Availability of data and materials

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Authors' contributions

SY and KT designed the study and developed the methodology. SY collected the data and performed the analysis with support from KT. Both the authors contributed to writing the manuscript. Both authors read and approved the final manuscript.

Competing interests
The authors declare that they have no competing interests.

Publisher's Note

Received: 12 May 2018 Accepted: 2 August 2018
Published online: 13 November 2018

References

Ashby, William J. 1982. The drift of French syntax. *Lingua* 57: 29–46.
Ashby, William J. 1988. The syntax, pragmatics, and sociolinguistics of left- and right-dislocations in French. *Lingua* 75: 203–229.
Baayen, R. Harald, Douglas J. Davidson, and Douglas M. Bates. 2008. Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language* 59: 390–412.
Barr, Dale J., Roger Levy, Christoph Scheepers, and Harry J. Tily. 2013. Random effects structure for confirmatory hypothesis testing: keep it maximal. *Journal of Memory and Language* 68: 255–278.
Bates, Elizabeth, and Brian MacWhinney. 1981. Second-language acquisition from a functionalist perspective: pragmatic, semantic, and perceptual strategies. *Annals of the New York Academy of Sciences* 379 (1): 190–214.
Branigan, Holly P., Martin J. Pickering, and Mikihiro Tanaka. 2008. Contributions of animacy to grammatical function assignment and word order during production. *Lingua* 118 (2): 172–189.
Caplan, David, Gayle DeDe, Gloria Waters, Jennifer Michaud, and Yorghos Tripodis. 2011. Effects of age, speed of processing, and working memory on comprehension of sentences with relative clauses. *Psychology and Aging* 26 (2): 439–450.
Chafe, Wallace. 1976. Givenness, contrastiveness, definiteness, subjects and topics. In *Subject and topic*, ed. Charles N. Li, 25–55. New York: Academic Press.
Chen, Po-ting. 2009. Topic and topic-comment structures in first language acquisition of Mandarin Chinese. In *Proceedings of the 21st North American Conference on Chinese Linguistics (NACCL-21)* 1: 165–177.
Coulmas, Florian. 2005. Communicating across generations: age as a factor of linguistic choice. In *Sociolinguistics: The study of speakers' choices*, 1st ed., 52–67. New York: Cambridge University Press.
Craik, Fergus I.M., and Ellen Bialystok. 2006. Cognition through the lifespan: mechanisms of change. *Trends in Cognitive Sciences* 10 (3): 131–138.
Ernst, Thomas, and Chengchi Wang. 1995. Object preposing in Mandarin Chinese. *Journal of East Asian Linguistics* 4 (3): 235–260.
Hou, John Yien-yao. 1979. *Grammatical relations in Chinese*. Doctoral dissertation. Los Angeles: University of Southern California.
Hsu, Yu-yin. 2008. The sentence-internal topic and focus in Chinese. *Annual Meeting of the Berkeley Linguistics Society* 34(1): 437–448.
Huang, James C.-T., Audrey Y.-H. Li, and Yafei Li. 2009. *The Syntax of Chinese*. Cambridge: Cambridge University Press.
Katz, Stacey L., and Carl S. Blyth. 2007. *Teaching French grammar in context: theory and practice*. New Haven: Yale University Press.
Kuno, Susumu. 1978. Japanese: A characteristic OV language. In *Syntactic typology: studies in the phenomenology of language*, ed. Winfred P. Lehmann, 57–138. Austin: University of Texas Press.
Kuznetsova, Alexandra, Per Bruun Brockhoff, and Rune Haubo Bojesen Christensen. 2016. lmerTest: tests in linear mixed effects models. https://cran.r-project.org/package=lmerTest.
Lee, Thomas Hun-tak. 2010. The acquisition of word order in a topic-prominent language: corpus findings and experimental investigation. In *Proceedings of the 24th Pacific Asia Conference on Language, Information and Computation*. http://www.aclweb.org/anthology/Y10-1003 .
Li, Jieqiong. 2010. *An experimental study of basic word order acquisition in Mandarin-speaking children*. Master's thesis. Changsha: Hunan University.
Li, Ping, Elizabeth Bates, Hua Liu, and Brian MacWhinney. 1992. Cues as functional constraints on sentence processing in Chinese. In *Language Processing in Chinese*, 207–234. Amsterdam: Elsevier.
Li, Ping, Elizabeth Bates, and Brian MacWhinney. 1993. Processing a language without inflections: a reaction time study of sentence interpretation in Chinese. *Journal of Memory and Language* 32: 169–192.
Lu, Hui-chuan. 1994. Second preverbal NPs in Chinese. In *Proceedings of the Sixth North American Conference on Chinese Linguistics (NACCL-6)*. Los Angeles: University of Southern California.
Meisel, Jurgen, Martin Elsig, and Esther Rinke. 2013. Language change across the lifespan. In *Language acquisition and change: a morphosyntactic perspective*, 20–51. Edinburgh: Edinburgh University Press.
Miao, Xiao-chun. 1981. Word order and semantic strategies in Chinese sentence comprehension. *International Journal of PsychoLinguistics* 8 (3): 109–122.
Miao, Xiao-chun, Guo-peng Chen, and Hou-chang Ying. 1984. Reexamination of the roles of word order and lexical meaning in Chinese sentence comprehension. *Psychological Science* 7(6): 1–7.
Oh, Se Jin, Jee Eun Sung, and Hyun Sub Sim. 2016. Age-related differences in animacy effects as a function of word-order canonicity in a verb-final language: evidence from ERP. *Communication Sciences & Disorders* 21 (4): 653–667.
Paul, Waltraud. 2002. Sentence-internal topics in Mandarin Chinese the case of object preposing. *Language and Linguistics* 3 (4): 695–714.
Paul, Waltraud. 2005. Low IP area and left periphery in Mandarin Chinese. *Recherches Linguistiques de Vincennes* 33: 111–133.
Paul, Waltraud. 2015. Sentence-internal topic vs sentence-external topic. In *New perspectives on Chinese syntax*, 239–240. Berlin: De Gruyter Mouton.
Qu, Yanfeng. 1994. *Object noun phrase dislocation in Mandarin Chinese*. Doctoral dissertation. Vancouver: University of British Columbia.
Shi, Yunzhi. 2001. The distinction between subject and topic in Chinese. *Linguistics Study* 43: 82–91.
Shyu, Shu-ing. 1995. *The syntax of focus and topic in Mandarin Chinese. Doctoral dissertation*. Los Angeles: University of Southern California.

Shyu, Shu-ing. 2001. Remarks on object movement in Mandarin SOV order. *Language and Linguistics* 2 (1): 93–124.
Stine-Morrow, Elizabeth, Sharon Ryan, and J. Sandra Leonard. 2000. Age differences in on-line syntactic processing. *Experimental Aging Research* 26 (4): 315–322.
Su, I-ru. 2001. Transfer of sentence processing strategies: a comparison of L2 learners of Chinese and English. *Applied PsychoLinguistics* 22 (1): 83–112.
Sun, Chaofen. 2006. *Chinese: a linguistic introduction*. New York: Cambridge University Press.
Sun, Chaofen, and Talmy Givón. 1985. On the so-called SOV word order in Mandarin Chinese: a quantified text study and its implications. *Language* 61 (2): 329–351.
Sung, Jee Eun. 2017. Age-related decline in case-marker processing and its relation to working memory capacity. *The Journals of Gerontology. Series B, Psychological Sciences and Social Sciences* 72 (5): 813–820.
Wagner, Suzanne Evans. 2012. Age grading in sociolinguistic theory. *Linguistics and Language Compass* 6 (6): 371–382.
Wagner, Suzanne Evans, and Gillian Sankoff. 2011. Age grading in the Montréal French inflected future. *Language Variation and Change* 23: 275–313.
Wang, Luming 王路明. 2015. A preference for word order or animate actor? An ERP study on processing ambiguous verb-final constructions in Chinese 优势语序还是优势解读?利用ERP考察汉语双论元歧义句的解歧过程. *Acta Psychologica Sinica 心理学报* 47 (7): 869–877.
Wang, Luming, Matthias Schlesewsky, Markus Philipp, and Ina Bornkessel-Schlesewsky. 2012. The role of animacy in online argument interpretation in Mandarin Chinese. In *Case, word order and prominence: interacting cues in language production and comprehension*, ed. Monique Lamers and Peter de Swart, 91–119. Dordrecht: Springer.
Wlotko, Edward W., Chia-lin Lee, and Kara D. Federmeier. 2010. Language of the aging brain: event-related potential studies of comprehension in older adults. *Language and Linguistic Compass* 4 (8): 623–638.
Xu, Nannan, Xiaopu Hou, Beibei Zhao, Zude Zhu, and Yiming Yang. 2017. Age-related temporal-spatial dynamic ERP changes during sentence comprehension. *Neuroscience Letters* 645: 74–79.
Zhu, Dexi 朱德熙. 1982. *Lectures on Chinese grammar 语法讲义*. Beijing: The Commercial Press.
Zhu, Zude, Xiaopu Hou, and Yiming Yang. 2018. Reduced syntactic processing efficiency in older adults during sentence comprehension. *Frontiers in Psychology* 9: 243 https://doi.org/10.3389/fpsyg.2018.00243.