The acceptability of two reflexives in one clause: identity, blocking and dialectual variations

Chaoyi Chen

[1] Rutgers University

The acceptability of two reflexives in one clause: identity, blocking and dialectual variations

Methods

Mandarin has two kinds of reflexives: the simple reflexive *ziji* 'self' and the complex reflexive *ta-ziji* 'him-self.' When appearing in a subordinate clause, both *ziji* and *ta-ziji* can refer to the subject in the matrix clause. For example, in the sentence '*John* thinks that *ziji/ta-ziji* is clever,' both *ziji* and *ta-ziji*' are able to refer to the matrix subject *John*. This paper examines the case where two reflexives appears in one subordinate clause and explores the factors affecting the acceptability of the sentence. In the sentence '*John* thinks that REF1 cheats REF2.' REF1 and REF2 can be occupied by either *ziji* or *ta-ziji*, thus four possibilities are generated: *ziji-ziji*, *ta-ziji-ta-ziji*, *ziji-ta-ziji* and *ta-ziji-ziji*. On top of these four cases, three factors of theoretical interests are considered in this study: (i) whether or not two reflexives are identical; (ii) whether or not a less specific reflexive precedes a more specific reflexive (*ziji-ta-ziji* vs the other three cases; I called the cases of a less specific reflexive preceding a more specific reflexive as **blocking** hereafter); (iii) whether or not the speaker is a northern Mandarin speaker. This is a hypothetical project and the data were stimulated by the R-package *faux*.

Participants

This study recruited a total of 100 participants, among which 50 participant were northern Mandarin speakers and 50 participants were southern Mandarin speakers (non-northern Mandarin speakers). Their northern/southern identities are defined in term of the dialects they speak. Participants were screened such that their age were greater than 18 years old.

Procedure

Participants completed the tests individually and finished the following tasks in order: a consent form (5 minutes), a language background questionnaire (5 minutes), an acceptability rating task (15 minutes).

Language background questionnaire

It was administered in the participants' L1 (Mandarin) and contained the questions about the dialect(s) they speak, the age of acquisition of Mandarin, knowledge of other languages, and language experience and use.

Materials

This study tested four sentences in total. These four sentences is framed as *Zhangsan* thinks that REF1 cheats REF2. The four sentences for test instantiate the four possibilities of reflexive combinations: *ziji-ziji*, *ta-ziji-ta-ziji*, *ziji-ta-ziji* and *ta-ziji-ziji*. For the identity variable, *ziji-ziji*, *ta-ziji-ta-ziji* are marked as yes and *ziji-ta-ziji* and *ta-ziji-ziji* are marked as no. For the blocking parameter, *ziji-ziji*, *ta-ziji-ta-ziji* and *ta-ziji-ziji* are marked as yes while *ziji-ta-ziji* is marked as no.

The acceptability rating task

This experiment was an acceptability rating task administered online using XXX, a data collecting website. Sentence items were presented one at a time which participants rated on a scale from 1 (very bad) to 7 (very good) intuitively as a formal measure of linguistic acceptability. These Likert scales have been widely used in psychological research (Likert 1932; Hartley 2014) as well as judgment replication studies. Participants were required to judge the acceptability of 20 sentences: 8 practice sentences and 12 non-practice sentences (8 fillers, 4 experimental). A one-minute break was placed after the first ten

sentences were rated. The practice sentences served to familiarize subjects with the task and 7-point rating system. The order of items presented in the experiment was randomized.

Results

Data summary

We obtained 400 observations in total from 100 participants and each participant rated on four test sentences. The sentence of four combinations *ziji-ta-ziji*, *ta-ziji-ziji*, *ziji-ziji* and *ta-ziji-ta-ziji* were marked as "s1,""s2,""s3" and "s4" respectively in the raw data and their identity and blocking properties were added to the data frame after the experiment (a part of data cleaning). The data frame for analysis is shown below. The participant from S001 to S050 are northern Mandarin speakers and the participant from S051-S100 are Mandarin Chinese speakers.

id	dia	sentence	rating	Blocking	Identity
S001	nor	s1	1	yes	no
S001	nor	s2	1	no	no
S001	nor	s3	4	no	yes
S001	nor	s4	5	no	yes
S002	nor	s1	2	yes	no
S002	nor	s2	3	no	no

Descriptive statistics

The descriptive statistics of acceptability ratings (mean, standard deviation(SD) and the number of observations) for the four sentences are given in the table below. The acceptability ratings with respect to blocking, identity and dialectual background are presented in Fig.1, Fig.2 and Fig.3 respectively.

Blocking	mean	SD	N	Identity	mean	SD	N
no	4.693	2.033	300	no	3.66	2.151	200
yes	3.38	2.206	100	yes	5.07	1.911	200

dia	mean	SD	N
nor	2.575	1.328	200
sou	6.155	1.037	200

The overall pattern suggest that the sentences without blocking (mean = 4.69; SD = 2.03) have higher scores than the sentences with blocking (mean = 3.38; SD = 2.21); the sentences with identical reflexives (mean = 5.07; SD = 1.91) have higher scores than the sentences with different reflexives (mean = 3.66; SD = 2.15); the acceptability rating of the sentences in question is lower for northern Mandarin speakers (mean = 2.58; SD = 1.33) than southern Mandarin speakers (mean = 6.16; SD = 1.04).
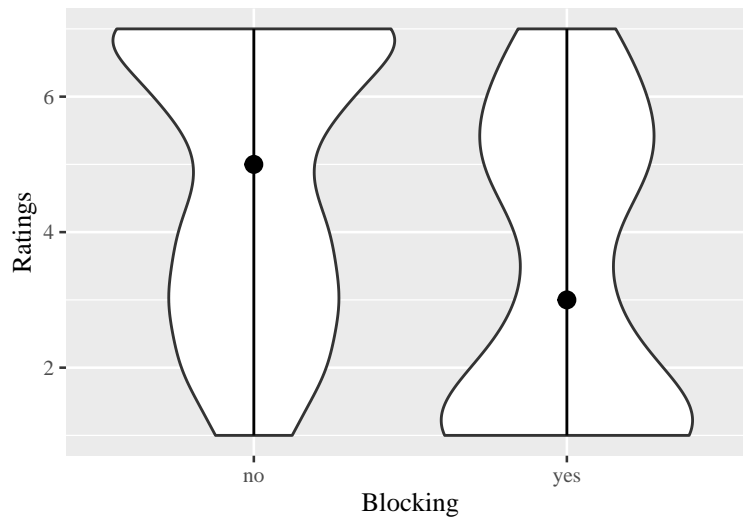

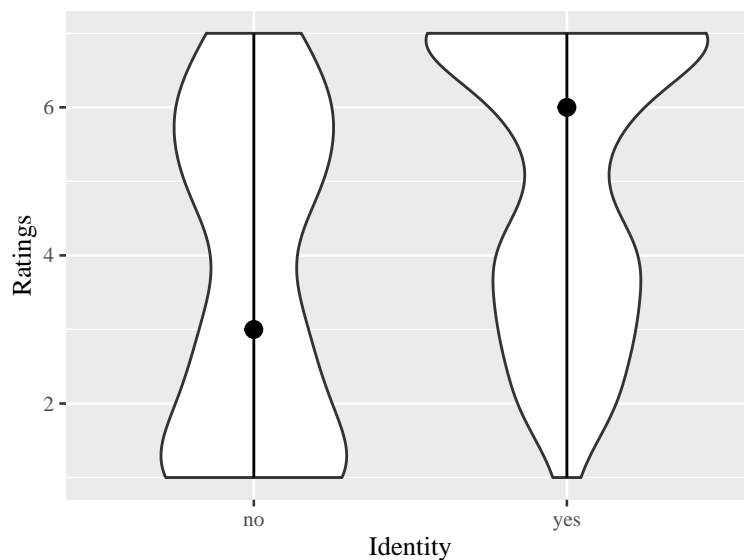*Figure 1*. The ratings of blocking and non-blocking sentences

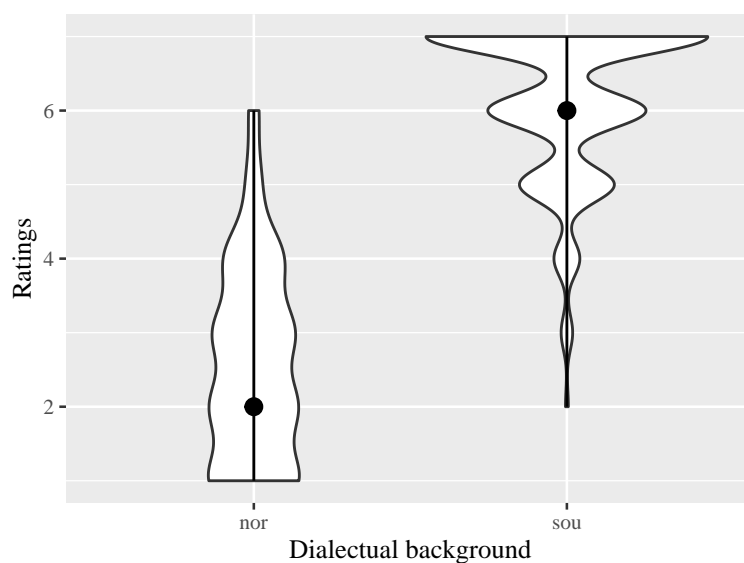*Figure 2*. The ratings of identical and non-identical sentences

Figure 3. The ratings from northern and northern participants

The ratings by individuals are presented in Fig 4. The participants S001-S050, who are northern Mandarin speakers (the left side of the table) have lower ratings in general than the participants S051-S100, who are southern Mandarin speakers (the right side of the table).
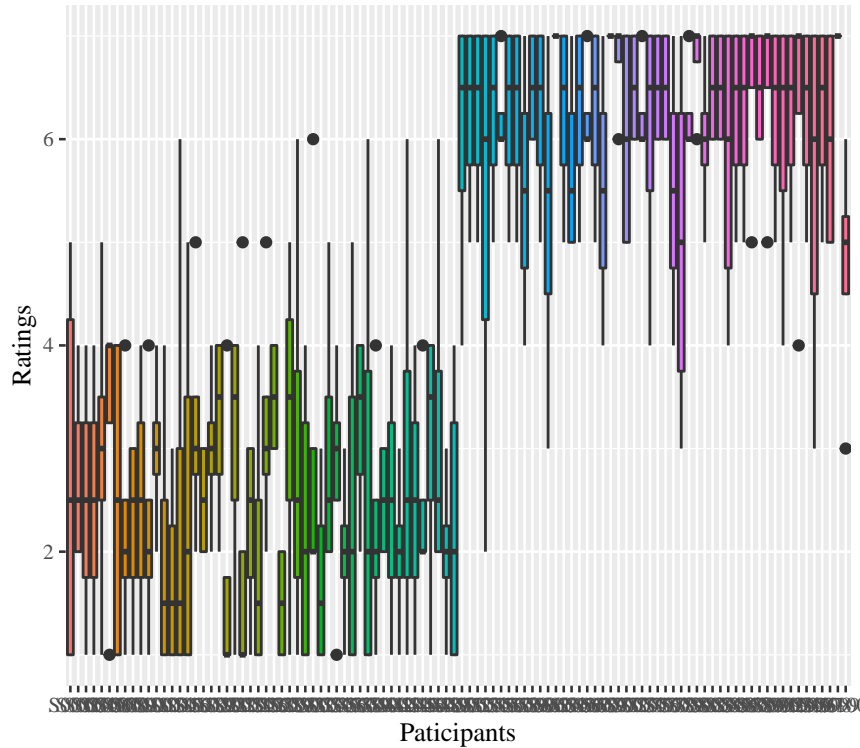
Figure 4. Individual Ratings from northern and southern participants

Cumulative link mixed model

Due to the ordinal nature of the outcome variable, this study used cumulative link mixed model (CLMM) to analyze the rating score data, and experiment-wise alpha was set at 0.05. The categorical fixed effect group was dummy coded with no (for Blocking), yes (for Identity) and northern speaker (for dialectual background) set as the baseline. The random effects structure only included by-subject random intercepts because the predictors are independent from each other. A total of five models were fitted: fm1 (only including intercept), fm2 (including the predictor Blocking), fm3 (including the predictor Blocking and Identity), fm4 (including the predictor Blocking, Identity and dialectual background) and fm5 (including the predictor Blocking, Identity and dialectual background as fixed effects and the random effects of subject). The comparison of these five models are given below.

	no.par	AIC	logLik	LR.stat	df	Pr(>Chisq)
fm1	6	1535	-762	NA	NA	NA
fm2	7	1505	-746	31.68	1	0.000
fm3	8	1486	-735	21.77	1	0.000
fm4	9	963	-473	524.18	1	0.000
fm5	10	964	-472	1.28	1	0.257

As a result, the model containing Blocking, Identity and dialectual background (fm4) had the lowest AIC value (1486) and thus provided the best fit of the data as shown in the table below. Blocking (z = -4.18, p < 0.001), Identity (z = 8.95, p < 0.001) and dialectual background (z = 16.76, p < 0.001) all have main effects on the final rating scores. Given the data from fm5, we are not confident enough to conclude that the random effect of participants show a main effect in this study (p = 0.257 > 0.05).

term	estimate	std.error	statistic	p.value	coef.type
1|2	-0.645	0.239	-2.70	0.007	intercept
2|3	1.002	0.235	4.26	0.000	intercept
3|4	2.471	0.267	9.25	0.000	intercept
4|5	3.987	0.327	12.19	0.000	intercept
5|6	5.477	0.389	14.10	0.000	intercept
6|7	7.371	0.455	16.21	0.000	intercept
Blockingyes	-1.138	0.272	-4.18	0.000	location
Identityyes	2.283	0.255	8.95	0.000	location
diasou	6.367	0.380	16.76	0.000	location

References

Hartley, J. (2014). Some thoughts on likert-type scales. *International Journal of Clinical and Health Psychology, 14*(1), 83–86.

Likert, R. (1932). A technique for the measurement of attitudes. *Archives of Psychology.*

R Core Team. (2021). *R: A language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/