Experimental evidence for the discourse potential of bare nouns in Mandarin*

Jess H.-K. Law & Kristen Syrett

Rutgers University

1. Introduction

Many studies on bare nouns have come to an agreement that these nominals have a special status that distinguishes them from regular indefinites, in the form of semantic (pseudo-)incorporation (Farkas & de Swart 2003, Chung & Ladusaw 2004, Dayal 2011, Krifka 2016, a.o.). However, there is far less consensus on whether (and if yes, how) this status is reflected in the domain of discourse. Chung & Ladusaw (2004) claim that incorporated nominals in Chamorro are transparent, i.e., they support anaphora just as regular indefinites do. However, Farkas & de Swart (2003), Dayal (2011), Modarresi (2014), and Krifka (2016) argue that bare nouns in Hungarian, Hindi and Persian have reduced discourse transparency as compared with indefinites.[1]

This study addresses the debate over the discourse status of bare nouns by presenting an experimental study of bare nouns in Mandarin. We compare pronominal anaphora with a bare noun, as exemplified in (1), to similar anaphora supported by a regular indefinite, as exemplified by (2), and ask to what extent bare nouns share the hallmark property of regular indefinites in supporting pronominal anaphora.[2]

(1) *Pronominal anaphora supported by a bare noun*
Wo kanjian-le **xuesheng**$_i$. *Ta/tamen*$_i$ gen wo wenhao.
I see-Asp student he/they with me greet
'I saw a student/students. He/they greeted me.'

*This work has benefited from discussions with Simon Charlow, Veneeta Dayal, and Ken Safir. Funding was generously provided by a startup grant to K. Syrett. Special thanks go to Judith Degen, Chelsea Sanker and Shu-hao Shih for tutorials on R and linear regression, and to the research assistants in Laboratory for Developmental Language Studies, who helped conduct the norming study.

[1]Reduced discourse transparency has been attributed to number-neutral bare nouns in Hungarian and Persian. It is not clear what role number morphology on bare nouns plays in Hindi. This paper only concerns the discourse property of number-neutral bare nouns.

[2]Bare nouns are number-neutral in Mandarin and have been suggested to be compatible with both singular and plural pronouns (Rullmann & You 2006).

(2) *Pronominal anaphora supported by an indefinite*
Wo kanjian-le **yi-ge xuesheng**$_i$. *Ta*$_i$ gen wo wenhao.
I see-Asp one-Cl student he with me greet
'I saw a student. He greeted me.'

To do so, we couple an offline, acceptability judgment task with an online, self-paced reading task (Just et al. 1982) to assess how native speakers of Mandarin judge pronominal anaphora with bare nouns and regular indefinites. Previewing our results, we find that the two types of pronominal anaphora are judged to have comparable acceptability. However, pronominal anaphora involving bare nouns incurs uncreased reading time, indicating additional processing. These results provide evidence in support of the claim that extra steps are required for bare nouns to support discourse anaphora (e.g., Dayal 2011, Krifka 2016).

The rest of the paper is organized as follows: Section 2 outlines the existing theories and their implications. Section 3 presents the experiment. We conclude in section 4.

2. Background

2.1 Three hypotheses

Previous literature offers three hypotheses concerning the discourse status of bare nouns, which we refer to as the transparency hypothesis, the opacity hypothesis, and the translucency hypothesis. In this section, I introduce these hypotheses and elaborate on how the current research tests the predictions arising from each of them.

Proponents of **the transparency hypothesis** take bare nouns to have similar discourse potentials as indefinites. For example, Chung & Ladusaw (2004) argue that bare nouns should be analyzed as semantically incorporated nominals, which they show to be able to support pronominal anaphora in Chamorro, as demonstrated in (3).

(3) *Chamorro incorporated nominal* (Chung & Ladusaw 2004:122)
Käda unu ni gäi-[**haga**$_i$], siempri ha-po'lu na bunita gui'$_i$
each one Comp WH[nom].have-daughter surely Agr-assume Comp pretty she
'Everyone who has a daughter$_i$ thinks that she$_i$ is beautiful.'

Other studies take bare nouns to have reduced discourse transparency. Within this camp, some studies take bare nouns to be totally opaque in discourse and other take them to be translucent, i.e., less transparent than indefinites but are still able to support pronominal anaphora.

The opacity hypothesis was first discussed by Dayal (1999) and later defended at great length by Farkas & de Swart (2003). According to Farkas & de Swart (2003), bare nouns in Hungarian cannot support pronominal anaphora, unlike regular indefinites, as evidenced by the contrast in (4) and (5) (examples cited from Farkas & de Swart 2003: (38–39)).

(4) *Hungarian pronominal anaphora with a bare noun*

a. János$_i$ **beteget**$_j$ vizsgált a rendelöben.
Janos patient.Acc examine.Past the office.in
'Janos patient-examined in the office.'

b. ??pro$_i$ Túl sulyosnak találta **öt**$_j$ és beutaltatta pro$_j$ a korházba.
pro too severe.Dat find.Past he.Acc and intern.Cause.Past pro the hospital.in
'He found him too sick and sent him to hospital.'

(5) *Hungarian pronominal anaphora with an indefinite*

a. János$_i$ **egy beteget**$_j$ vizsgált a rendelöben.
Janos a patient.Acc examine.Past the office.in
'Janos examined a patient in the office.'

b. pro$_i$ Túl sulyosnak találta **öt**$_j$ és beutaltatta **pro**$_j$ a korházba.
pro too severe.Dat find.Past he.Acc and intern.Cause.Past pro the hospital.in
'He found him too sick and sent him to hospital.'

The degraded status of (4) is attributed to the interaction between the bare noun and the pronoun: the pronoun carries a presupposition that a discourse referent (introduced by the bare noun antecedent) is present in the context. Since the bare noun fails to introduce such a discourse referent, the presupposition cannot be met. As a result, presupposition failure kicks in and pronominal anaphora with a bare noun is severely degraded.

Lying somewhere between these two extremes is what we termed **the translucency hypothesis**, of which Dayal (2011), Modarresi (2014), and Krifka (2016) are representative proponents.[3] This hypothesis takes pronominal anaphora with bare nouns to *appear* transparent. However, the transparency of bare nouns crucially *differ* from the transparency of regular indefinites, in that bare nouns do not introduce discourse referents the way that regular indefinites do. Nonetheless, pronominal anaphora with bare nouns can be achieved via adjustments in *pragmatics*, which are not required for pronominal anaphora with regular indefinites.

This hypothesis has two variants. In the first variant, Dayal (2011) and Krifka (2016) suggest that bare nouns have reduced discourse transparency in the sense that they do not introduce discourse referents. However, by making (potentially effortful) indirect reference to the event that has a bare noun as its thematic participant, pronominal anaphora may still be supported. In the second variant, Modarresi (2014) argues that bare nouns do introduce discourse referents, but these discourse referents are number-neutral (see also Kamp & Reyle 1993). Since overt pronouns typically bear a singular or plural number presupposition, this presupposition has to be accommodated when the relevant discourse referent is number-neutral.

[3]Farkas & de Swart (2003) first introduced this terminology to refer to *bare plurals* in Hungarian, which they take to be able to support pronominal anaphora with extra adjustments. We have extended the terminology to cover bare nouns in general that may require additional steps to support pronominal anaphora.

2.2 Predictions of the hypotheses

Assuming that an offline acceptability judgment task tracks acceptability and an online reading task tracks processing effort as a function of the reading time, these three hypotheses give rise to three distinct predictions regarding the behavioral responses in the two types of experimental tasks, as shown in (6). If bare nouns are *transparent*, as claimed by Chung & Ladusaw (2004), anaphora involving them should be judged acceptable in the judgment task and associated with fast response time (RT) in the reading task, just like the indefinite baseline would. If, on the other hand, bare nouns are *opaque*, as suggested by Farkas & de Swart (2003), they should differ from the baseline in both tasks: anaphora involving them should be judged unacceptable in the judgment task and associated with slower response time in the reading task.

Alternatively, if bare nouns are *translucent*, we should expect anaphora with coordinated and uncoordinated bare nouns to pattern like the indefinite baseline in the judgment task. However, the two types of bare nouns may or may not contrast in the processing task, depending on the reason of the translucency. If bare nouns are translucent due to number-neutrality (Translucency 1), as claimed by Modarresi (2014), then we should expect anaphora with coordinated bare noun antecedents to be processed faster than that with uncoordinated bare noun antecedents. This is because number-neutrality is lost with coordinated bare nouns. However, if translucency is due to their inability to introduce discourse referents (Translucency 2), as argued by Dayal (2014) and Krifka (2016), then we should expect anaphora with coordinated and uncoordinated bare nouns to pattern alike and contrast with anaphora with indefinites in the processing task.

(6) Predictions from the three hypotheses regarding anaphora with bare nouns

Hypothesis	Judgment task	Reading task	
Transparency	Acceptable	Fast RT	
Opacity	Unacceptable	Slow RT	
		Uncoordinated BNs	Coordinated BNs
Translucency 1	Acceptable	Slow RT	Fast RT
Translucency 2	Acceptable	Slow RT	Slow RT

3. Experiment

3.1 Stimuli

The stimuli consisted of 36 targets, 9 controls, and 65 fillers. Each target and control stimulus consisted of two sentences, as exemplified in (7). The first sentence (S1) introduced an antecedent noun phrase. The second sentence (S2) began with a (possibly covert) pronoun that was unambiguously anaphoric to the noun phrase introduced in S1.

(7) Sample stimulus

S1: Women zai bianlidian kandao-le xiaotou$_i$.
we at convenience.store saw-Asp thief
'We saw a thief/thieves in a convenience store.'

S2: {Ta/tamen/pro}$_i$ tou-wan dongxi zhihou jiu likai-le xianchang.
he/they/pro steal-Asp thing after then leave-Asp scene
'He/they/pro left after stealing things.'

All the S2s shared the same schema, as shown in (8). Each S2 could be divided into 6 fragments, including a pronoun region and the spill-over region. Previous studies have demonstrated that a spill-over region of roughly this size was appropriate for self-paced reading studies on pronouns (e.g., Chow et al. 2014). Each target fragment was additionally controlled for length, log word frequency, and visual complexity. The pronoun+1 region was controlled to be a verb but the parts of speech of other regions were not controlled. A linear-mixed effects model was fit to the data and it revealed that parts of speech did not have a main effect on the reading speed. In the self-paced reading task, all sentences were displayed fragment by fragment in a moving window paradigm.

(8) Sample stimulus format

S2	Ta/tamen/pro	tou-wan	dongxi	zhihou	jiu	likai-le.
Region	pronoun	pronoun+1	pronoun+2	pronoun+3	pronoun+4	pronoun+5
		spill-over region				

The target stimuli involved two factors, namely, bare noun type (with four levels) and pronoun type (with three levels). Crossing the two factors resulted in 12 types of target stimuli. Three sentences with different bare nouns used in different contexts were created for each type of target stimuli, giving rise to 36 target stimuli in total.

The four levels in bare noun type were: (i) uncoordinated bare nouns with a neutral number bias (BN-Neutral), (ii) uncoordinated bare nouns with a singular number bias (BN-Singular), (iii) uncoordinated bare nouns with a plural number bias (BN-Plural), and (iv) coordinated bare nouns. Bare nouns with different number biases were included because previous studies had shown that number bias was a factor relevant for pronominal anaphora with bare nouns (Modarresi 2014). More specifically, it was found that number-neutral bare nouns may have a singular or plural contextual number bias, and hence may be more compatible with a singular or plural pronoun. The number bias of the bare nouns used in this study was determined by an independent norming study. Coordinated bare nouns were included to test whether number-neutrality is the reason for discourse translucency of bare nouns.

The three levels in pronoun type are: (i) overt 3rd-person singular pronoun (SG), (ii) overt 3rd-person plural pronoun (PL), and (iii) covert pronoun.[4]

[4]Stimuli involving the covert pronoun turned out to have both low acceptability and slow RT, regardless of whether the antecedent was a bare noun or an indefinite. As it is likely that there are other complexities involved with processing the covert pronoun, these items were excluded from the analysis.

All control stimuli have numeral phrases as the referential antecedent. Numeral phrases, qua regular indefinites in Mandarin, were included to tease apart the transparency, opacity and translucency hypotheses. By comparing the behavior of pronominal anaphora with bare nouns to that of regular indefinites, we can determine which of these hypotheses best accords with the experimental results.

Stimuli were grouped into three lists using a Latin-Square design with pseudo-randomization, and were presented in Simplified Chinese characters (font size 28). Each list had 12 targets, 3 controls, and 31 to 33 fillers. The same targets and controls showed up in the self-paced reading task and the acceptability judgment task. A subset of the fillers used in the self-paced reading task were also used in the acceptability judgment task.

3.2 Participants and procedure

30 native Mandarin speakers were recruited to participate in the experiment. They were compensated with a seven-dollar bookstore gift card. All speakers participated in the acceptability judgement task and the self-paced reading task.

The experiment took place in a quiet laboratory setting, with two tasks presented back to back. The self-paced reading task was conducted before the acceptability judgment task to avoid familiarity effects. The experiment was conducted using a 21" computer screen with SuperLab 5. Instructions and practice trials were presented before each task.

In the self-paced reading task, participants read each item fragment by fragment (as presented in (8)), pressing the space bar to advance through the sentence pair. The reaction time between key presses were recorded. Comprehension questions appeared after some number of the stimuli, to ensure that participants were attentive and correctly interpreted the co-reference relation intended by the pronoun in each item. Data points with incorrectly answered comprehension questions were excluded in data analysis.

After a self-timed break of approximately one to two minutes, participants proceeded to the acceptability judgment task. In each trial, participants had to rate, based on a scale of 1 to 5, how acceptable the sentences presented to them (like the one in (7)) were by pressing the number keys on a keyboard to indicate their choice. The Chinese equivalents of the following descriptive labels were used with the scale: 5: Perfectly acceptable; 4: Acceptable; 3: Marginally acceptable; 2: Hardly acceptable; and 1: Completely unacceptable.

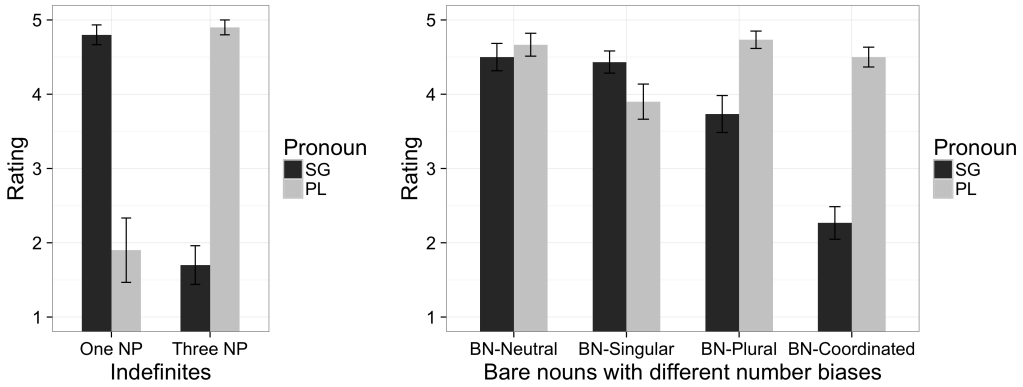
3.3 Results

All statistical analyses reported in this paper were conducted by linear mixed-effects modeling with lme4 package for the statistical language R (R Core Team 2016). All models used to generate the statistical analyses in this study are mixed-effects models with subject as the random effect. Models are fit using a restricted maximal likelihood technique. Probabilities were estimated by means of the function summary in the package lmerTest.

3.3.1 Acceptability judgment

A total of 450 ratings were collected from the target stimuli in the acceptability judgment task. No outliers were identified. The results are presented in (9), with the left figure showing the results from the control condition, and the right figure showing the results from the target condition.[5]

(9) Acceptability of pronominal anaphora with indefinites (left) and bare nouns (right)



As shown in the figure on the left in (9), the control items generated expected acceptability ratings. In particular, participants judged as unacceptable when a pronoun and an indefinite did not match in number. In addition, they rated as highly acceptable pronominal anaphora with indefinites (mean = 4.85). Thus, participants performed as expected with these control items in demonstrating that they used the full range of the scale.

The results from the target stimuli are presented in the figure on the right. There are a few points to note about this graph. First, in contrast to the sensitivity to different pronouns shown by indefinites, uncoordinated bare nouns had high acceptability regardless of the pronoun. Relatedly, neutral bare nouns without a number bias can support a singular pronoun and a plural pronoun equally well. These results are expected as bare nouns are number-neutral in Mandarin. That being said, number bias does show its effect-bare nouns with a particular number bias prefers a pronoun with the same number.

Second, focusing on bare nouns supporting pronouns that are compatible with their number bias, we found that they were also judged highly acceptable (mean = 4.57). Although the target stimuli had a slightly lower rating than the control stimuli when only considering the cases with a matching pronoun, the difference is not statistically significant ($\beta = 0.25$, SE = 0.19, $t = 1.31$, $p = 0.192$). What these high acceptability ratings indicate is that bare nouns are either transparent or translucent in discourse, but not opaque. The re-

[5]In addition to excluding data points from stimuli involving the covert pronoun, a set of data involving an indefinite in the form of a vague numeral phrase *ji-ge* NP 'several NP' were also excluded. The reason is that participants gave unexpected tolerance to pronominal anaphora with a singular pronoun. We speculate that this phrase may have been mistaken as an interrogative word, which has the same form and roughly means 'how many' in Mandarin.

sults from the self-paced reading task, to be discussed in the next subsection, are intended to tease part the transparency hypothesis and the translucency hypothesis.

3.3.2 Self-paced reading

Overall the comprehension questions received a high accuracy rate (mean = 95.38%), indicating that the participants paid attention in the reading task and processed the material as intended. In addition, there is a negative correlation between reaction time and acceptability rating in the data collected (r = -0.33, n = 45, df = 43, $p < 0.05$), confirming the assumption that participants generally spend more time reading less acceptable sentences.

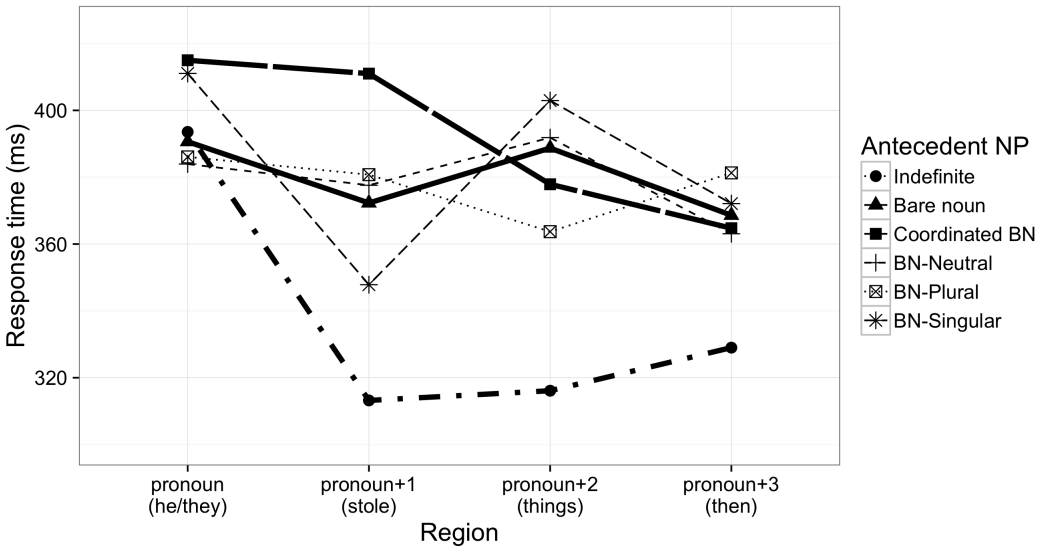
A total of 14292 response time data points were collected, 2550 of which were from the targets and controls. Response times larger than the upper quartile plus three times the interquartile range (1092ms) were taken to be outliers and removed from analysis. In addition, a lower cutoff point was set at 200ms because previous studies have established that any time shorter than this is likely due to unintended button presses rather than responses after meaningful semantic processing (e.g., Jegerski 2014). All response time data from items whose corresponding comprehension questions were incorrectly answered were removed. This series of data treatment affected 9.29% of the target and control data points, leaving 2313 target data points to be analyzed. The response time data were then log-transformed to make the distribution less skewed and better fit the statistical analyses.

The result (pre-log-transformation) is presented in (10). Let us first focus on the indefinite and the uncoordinated bare noun average, represented by the (thickened) dash-and-dot line decorated with circles at the bottom of the graph and the (thickened) solid line decorated with triangles in the middle of the graph, respectively. In the pronoun region, they do not differ (394ms vs. 391ms). However, they start to diverge in the pronoun+1 region (313ms vs. 372ms), and continue to do so in the pronoun+2 region (316ms vs. 389ms) and the pronoun+3 region (329ms vs. 369ms). The overall difference is statistically significant ($\beta = 0.12$, SE = 0.04, $t = -2.99$, $p < 0.05$).

The bare noun average is further decomposed into three (non-thickened) lines, the BN-Neutral line, the BN-Plural line, and the BN-Singular line, to show the internal composition. As shown in the graph, these three lines have a similar profile that is distinct from the indefinite's, indicating that whatever triggers an increased processing time was shared by all the bare nouns, regardless of their number bias.

Due to the lag time in processing, a slow down is often not seen in the 'trigger', but reflected on the spill-over region of that trigger. In this case, the pronoun does trigger a slow down in the spill-over region when its antecedent is a bare noun, we can then conclude that there is increased processing effort associated with pronominal anaphora supported by bare nouns. This result, when interpreted independently from the result of the acceptability judgment task, is compatible with a view that bare nouns have reduced discourse transparency, or are either translucent or opaque. However, in conjunction with the results of the acceptability judgment task, the combined results support the translucency hypothesis.

(10) Results from the self-paced reading task



Finally, the processing profile of pronominal anaphora with a coordinated bare noun antecedent is represented by the long-dash line decorated with squares. It is slight elevated in the first two regions compared to the uncoordinated bare noun average but the difference between them is not statistically significant ($\beta = 0.03$, SE = 0.04, $t = 0.82$, $p = 0.42$). However, the difference between the coordinated bare noun average and the indefinite average is statistically significant ($\beta = -0.12$, SE = 0.05, $t = -2.62$, $p < 0.05$). What these results suggest is that bare nouns are translucent not due to the number-neutrality of the discourse referents they introduce, but due to other factors. This finding favors the translucency hypothesis of Dayal (2011) and Krifka (2016) over that of Modarresi (2014).

3.3.3 Discussion

We started out with three hypotheses regarding the discourse transparency of bare nouns in Mandarin: the opacity hypothesis, the transparency hypothesis, and the translucency hypothesis. The acceptability judgment task showed that bare nouns support pronominal anaphora just as well as indefinites, refuting the opacity hypothesis. The self-paced reading task revealed that there was more processing effort associated with pronominal anaphora with a bare noun antecedent, as compared to pronominal anaphora supported by indefinites, refuting the transparency hypothesis. Moreover, since there is no contrast between coordinated and uncoordinated bare nouns in inducing increased processing effort, the version of the translucency hypothesis that attributes the increased processing to the number-neutrality of the discourse referent introduced by a bare noun is refuted. What is left then, is the translucency hypothesis that claims that bare nouns have inherently reduced discourse transparency due to their inability to introduce discourse referents; anaphora is still possible, as reflected by the high acceptability, however, extra pragmatic steps are required to achieve this result, as reflected by the increased processing time.

4. Conclusion

The findings of this study contribute to the literature on the discourse property of bare nouns at two important levels. At the methodological level, we present the first study to investigate the discourse property of bare nouns in an experimental setting, with tasks that are sensitive to offline acceptability judgments and real-time processing. Thus, at the theoretical level, the combined findings reveal what it means for bare nouns to have reduced discourse transparency, and be classified as "discourse translucent": they do support pronominal anaphora, but doing so incurs a processing effort in a way that pronominal anaphora with indefinites does not.

References

Chow, Wing-Yee, Shevaun Lewis, & Colin Phillips. 2014. Immediate sensitivity to structural constraints in pronoun resolution. *Front Psychol* 5:630.

Chung, Sandra, & William A. Ladusaw. 2004. *Restriction and saturation*. MIT Press.

Dayal, Veneeta. 1999. Bare NP's, reference to kinds, and incorporation. In *Proceedings of Semantics and Linguistic Theory 9 (SALT-9)*.

Dayal, Veneeta. 2011. Hindi pseudo-incorporation. *Natural Language and Linguistic Theory* 29:123–167.

Dayal, Veneeta. 2014. Bangla plural classifiers. *Language and Linguistics* 15:47–87.

Farkas, Donka, & Henriëtte de Swart. 2003. *The semantics of incorporation: From argument structure to discourse transparency*. CSLI Publications.

Jegerski, Jill. 2014. Self-paced reading. In *Research Methods in Second Language Psycholinguistics*, ed. Jill Jegerski & Bill VanPatten, 20–49. Routledge.

Just, M. A., P. A. Carpenter, & J. D. Wooly. 1982. Paradigms and processes in reading comprehensions. *Journal of Experimental Psychology: General* 111:228–238.

Kamp, Hans, & Uwe Reyle. 1993. *From discourse to logic*. Dordrecht: Kluwer Academic Publishers.

Krifka, Fereshteh, Manfred abd Modarresi. 2016. Number neutrality and anaphoric update of pseudo-incorporated nominals in persian (and weak definites in english). In *Proceedings of Semantics and Linguistic Theory 26*, 874–891.

Modarresi, Fereshteh. 2014. Bare nouns in persian: Interpretation, grammar and prosody. Phd. dissertation, University of Ottawa, Ottawa, Canada.

Rullmann, Hotze, & Aili You. 2006. General number and the semantics and pragmatics of indefinite bare nouns in Mandarin Chinese. In *Where semantics meets pragmatics*, ed. Klaus von Heusinger & Ken P. Turner, 175–196. Amsterdam: Elsevier.

Jess H.-K. Law, Kristen Syrett
jesshk.law@rutgers.edu, kristen.syrett@rutgers.edu