

The Hong Kong Polytechnic University

The Hong Kong Polytechnic University

Determining the types of contrasts: the influences of prosody on pragmatic inferences in Mandarin

This study explores the issues involving pragmatic inferences without either syntax or semantics cues. Although there is a rich and well-established literature from multiple languages demonstrating how pragmatic inferences can be triggered by syntactic and semantic cues, few studies focused on how pragmatic inference is triggered by cues outside of syntax and semantics. In Mandarin Chinese, the same sentence containing a numeral-classifier phrase as a negative polarity item can be employed for two types of scalar inferences based on either the numeral or the noun. The sentence *wo yi zhi ma 'i dou mei kan dao* ("I didn't even see one ant") can induce two different scalar inferences: 'I did not see one ant, much less two ants, three ants, and so on' by contrasting minimal quantity of one; or 'I did not see an ant, much less a dog, a cat, a human being, and so on' by contrasting the minimal size of ants. Since the sentence has the same syntactic structure and lexical items, Mandarin native speakers have to rely on other cues to construct appropriate pragmatic inferences in contexts. The three experiments in the study suggest that Mandarin speakers can use sentence prosody to determine which inference is intended, at least in experimental contexts that directly probe explicit awareness of prosody. Each prosodic pattern evokes a specific set of scalar inferences. The results suggest that prosodic prominence can serve as linguistic cues to pragmatic inferences.

Keywords: prosody; scalar inferences; numeral-classifier phrases; negative polarity items

Introduction

Pragmatics investigates the relation of signs to their users and interpreters [1]. The studies of pragmatics focus on the context-dependent meanings which are systematically abstracted from the logical form or the content of a construction

1 concerned in syntax and semantics [2, 3]. In order to interpret the information from the
 2 speaker, the hearer has to take the interaction of grammatical structure and context into
 3 consideration. The scalar inferences discussed in this study are the cases showing that
 4 the hearer evokes a mental scalar model from a grammatical construction and context
 5 [4].

6 The study investigates whether prosody influences pragmatic inferences by
 7 examining the types of inferences inferred from negative polarity items (NPIs). As a
 8 tonal language, Chinese has both syllable-level lexical tones and sentence-level
 9 intonation. The syllable-level lexical tones have been described as "small ripples riding
 10 on large waves of intonation" [5]. Intonation interacts with syllabic tones without
 11 cancelling their acoustic effects. The prominence of intonation is regarded as expanding
 12 pitch range. For example, the targets have larger pitch range, longer duration, and
 13 higher intensity in prosody [6]. Particularly, the contextually focused words in a
 14 sentence are prominent in pitch height and intensity [7]. The present study examines
 15 how sentence-level intonation, particularly focus, influences the interpretation of
 16 negative polarity items (NPIs).

17 NPIs are expressions which are only grammatical under certain semantic
 18 contexts, such as negation [8, 9]: for example, in English, *I haven't ever been to France*
 19 is grammatical but **I have ever been to France* is not. NPIs have been observed across
 20 many languages [10]. They are often words referring to very small amounts, e.g., *I*
 21 *didn't sleep a wink, He won't spend a red cent, They don't give a rat's ass about this*
 22 *topic*. In such cases, the negation of such a small amount allows the hearer to infer that
 23 larger amounts are also not true: e.g., if somebody did not sleep "a wink" then they
 24 surely did not sleep for a long time either. These types of small-quantity expressions
 25 which occur in environments related to negation are called *minimizers*, and are a type of

1 NPI. Across languages, minimizers are widely employed for pragmatic emphasis, due to
 2 their robust scalar inferences [8, 9]. Minimizers induce scalar reasoning because they
 3 evoke a mental scalar model with all the alternatives ranked for contrasting [11]. Since
 4 minimizers refer to an endpoint of a scale, they can contrast with all the other
 5 alternatives along the scale for emphasis [12, 13].

6 In numeral-classifier languages, ‘one’-phrases, which are composed of the
 7 numeral ‘one’, a classifier or a measure word, and a noun, are pervasively used as
 8 minimizers, as in example (1) below. Specifically, individual classifiers are employed
 9 for categorizing a semantically salient perceptual property of a noun which can be
 10 individuated [14].

11 (1) wo yi zhi changying dou mei kandao
 12 I one CLF fly FOC NEG see

13 ‘I did not see any flies.’ Just as in the examples above, sentences with numeral-
 14 classifier phrases like (1) also elicit inferences about what the phrase is being contrasted
 15 with. (In this and other examples, CLF stands for classifiers, FOC for focus markers, and
 16 NEG for negation.) Specifically, for a sentence like (1), two types of inferences are
 17 possible. The sentence can infer that the speaker saw ‘not even one fly, *much less two*’
 18 if the minimizer is interpreted as invoking a quantity-based contrast, while it can instead
 19 imply that the speaker saw ‘not even one fly, *much less one human being*’ if the
 20 minimizer is interpreted as invoking a type-based contrast. In the quantity-contrast
 21 interpretation, the minimal amount that is being invoked is "one", and this is raised in
 22 contrast with greater amounts ("two flies", "three flies", etc.); in the type-contrast
 23 interpretation, the minimal amount is some type of noun that has a high probability of
 24 occurring in this context. For example, this sentence is uttered in a context where there
 25 are likely to be flies, and this is raised in contrast with nouns that are even less likely or

1 prototypical in this context. The quantity-contrast interpretation is straightforward due
2 to the involvement of a numeral phrase, while the type-contrast interpretation is
3 relatively less straightforward since it is relevant to the setting of the contexts. However,
4 it is clear that the noun chosen for contrasting is the proposition which is assumed to be
5 the most likely one.

6 In other numeral classifier languages such as Japanese and Korean, the distinction
7 of the two types of inferences is reflected in morphology and word order [15, 16].
8 However, in Mandarin, the two sets of inferences occur in the same word order, syntax,
9 and semantics. Native Mandarin speakers thus require other cues to discern the
10 pragmatic differences. It has been noted in studies of NPIs that minimizers are claimed
11 to tend to occur in constructions which can attract people's focus [9]. For instance, an
12 expression interpreted as a minimizer carries an emphasized intonation which is
13 different from its other uses. In line with this observation, Mandarin minimizers tend to
14 occur in the preverbal construction as in (1): this sentence has a Subject-Object-Verb
15 word order, which differs from the Subject-Verb-Object word order that is canonical
16 and unmarked in Mandarin. This preverbal position, where "one fly" occurs in sentence
17 (1), has received substantial attention in the literature and has been regarded to carry
18 focus [17, 18, 19]. It is also noted that 'one'-phrases may bear a different prosodic stress
19 when they are used as minimizers as opposed to when they are used normally [20].
20 According to these studies, a connection between prosodic stress, focus of attention, and
21 pragmatics can be inferred. However, the issues of how focus is perceived by native
22 speakers and of whether prosodic stress modulates the inferences drawn by speakers in
23 this type of sentence have been barely touched upon. Thus, the experiments of this
24 study are designed to answer these questions by investigating whether Mandarin
25 speakers are aware of the association between prosody and pragmatics.

1 The critical stimuli of the three experiments in this study are sentences with the
 2 structure exemplified in (2). A prosodic stress is superimposed either on the numeral-
 3 classifier constituent or on the noun of the numeral phrase. As in (2), the same sentence
 4 has different highlights in prosody, as shown in the bolded sections. The stimuli were
 5 produced by a female native Mandarin speaker, who speaks only Beijing Mandarin
 6 without other dialects.

7 (2) a. jintian maomi kafeiguan mei kai, **yi zhi** maomi dou mei you

8 b. jintian maomi kafeiguan mei kai, yi zhi **maomi** dou mei you

9 today cat café NEG open oneCLF cat FOC NEG exist

10 ‘The cat café is closed today. I didn’t see even one cat.’

11 Although other numeral-classifier languages such as Japanese can rely on
 12 morphology to distinguish the two types of scalar inferences, it has been noted that the
 13 elements attached by a scalar particle, such as the noun or the numeral-classifier unit of
 14 a numeral phrase, carry an emphatic prosody [16]. In the setting of a quantity contrast,
 15 the numeral-classifier unit is stressed; in the setting of a type contrast, the noun is
 16 stressed. Therefore, our intuition suggests that the prosody in (2a) should be more likely
 17 to evoke a quantity contrast (i.e., an interpretation like "I didn't even see *one* cat, let
 18 alone two cats, three cats, etc."), whereas the prosody in (2b) should be more likely to
 19 evoke a type contrast (i.e., an interpretation like "I didn't even see one *cat*, let alone one
 20 person, one bird, etc.").

21 Each experiment has a different task for the participants to respond to the stimuli.
 22 In Experiment 1, the participants were asked to judge whether the sentence which they
 23 heard from an audio clip was consistent with a paragraph they read previously, which

1 set up a context consistent with either a quantity contrast or a type contrast. The design
2 of Experiment 2 is the same as that of Experiment 1, but the participants were asked to
3 give consistency ratings on a Likert scale rather than binary judgments of consistency.
4 In Experiment 3, the participants read a context and then heard two auditory versions of
5 the sentence with different prosody, and were instructed to select the version that better
6 fit the context. The three tasks were made to test whether prosody is a determinant of
7 the types of scalar reasoning and how much Mandarin speakers are aware of prosody.
8 The results can help to validate the associations of the unconnected pieces in the
9 literature of focus, prosody, and pragmatic inferences.

10 In each experiment, all the participants provided their informed consent before
11 they began the survey. The experimental procedures were approved by the Hong Kong
12 Polytechnic University. Each experiment had both a traditional character version and a
13 simplified character version. The traditional character version was distributed in Taiwan
14 and Hong Kong, while the simplified character version was distributed in mainland
15 China. When the survey was advertised through the platforms of social media, both
16 links were provided and volunteers could choose based on their preference.

17 **Experiment 1**

18 *Methods*

19 *Participants*

20 Sixty-nine native speakers of Mandarin, 60 users of traditional Chinese characters and 9
21 users of simplified Chinese characters, were included in the first experiment. Two were
22 removed from analysis because they did not correctly respond to baseline questions (see
23 *Procedure*), leaving 67 participants (aged 20-60, mean 31) in the final analysis.

1 **Materials**

2 The experiment stimuli comprised 12 sentences along with 16 fillers. A short paragraph
 3 is provided to set up the relevant context for each stimuli sentence, as shown in (3). The
 4 sentence always referred to some set that did not have some property. As in (3), the dog
 5 park does not have dogs, which should be expected to be most likely encountered in the
 6 defined setting. Another type of noun, human being, is involved as an alternative to be
 7 contrasted with dogs in this setting. (3)a includes the presence of human being, whereas
 8 (3) specifies no presence of human being.

9 (3) *Zhangwei dao le youmingde liugou gongyuan, pingchang zheli henduo*
 10 *gou.* ‘Zhangwei went to a famous dog park. Usually there were a lot of
 11 dogs.’

12 a. *Jintian meiyou gou que you ren zai gongyuan li sanbu* ‘Today there were no
 13 dogs, but there were people walking in the park.’ [Yes-context]

14 b. *Jintian gongyuan li meiyou gou ye meiyou ren* ‘Today there were neither dogs
 15 nor people in the park’ [No-context]

16 *Zhangwei huilei hou gen ni shou:* ‘Zhangwei came back and told you:’

17 The experimental stimuli appeared in the syntactic format of (1). In the case of (3), the
 18 contents of the audio files are presented in (4).

19 (4) *Wo jiantian qu le liugou gongyuan,* ‘I went to the dog park today.’

20 a. **yi zhi** gou dou mei kandao

21 one CLF dog FOC NEG see

22 b. yi zhi **gou** dou mei kandao

23 one CLF dog FOC NEG see

24 ‘I did not see even one dog.’

1 The two versions of the audio files both state the lack of a specific property which is the
 2 most expected in the defined set. The only difference is that one has the prosodic stress
 3 on the noun, while the other stresses the numeral-classifier combination in the prosody.
 4 For ease of reference, the former is termed noun stress, while the latter is called
 5 quantifier stress. Based on the four conditions, the experiment followed a 2×2 design:
 6 PROSODIC STRESS (*noun stress* vs. *quantifier stress*) × CONTEXT (*type alternative*
 7 *present* vs. *type alternative absent*). The items were organized into four lists in a Latin
 8 square design.

9 The fillers can be divided into two groups. The first group contains six sentences
 10 which mismatches the content from the audio files. There are three types of mismatches
 11 including number, quantity, and location. One example of number mismatch is provided
 12 in (5), where the semantics of the two contexts are completely inconsistent. The fillers
 13 both serve as a check that the received data are valid, and to distract participants from
 14 the experimental manipulation.

15 (5) Reading context: *Mama qie le san ge pingguo, danshi meiyong chi. Baba*
 16 *gen ni shuo*: ‘Mom cut three apples, but didn’t eat them. Dad told you:’
 17 Audio context: *Mama yi ge pingguo dou mei qie*. ‘Mom did not cut even one
 18 single apple.’

19 The other group of fillers includes ten sentences from another experiment for
 20 investigating the scalar implicatures from Mandarin *youxie* ‘some’. The full list of
 21 stimuli is available at <https://osf.io/nsgfv/>.

22

23 ***Procedure***

1 This experiment was administered online via Ibex Farm (<http://spellout.net/ibexfarm/>).
 2 At the beginning of the experiment, participants indicated their consent to participate,
 3 provided demographic information (age, sex, and native language), and answered two
 4 questions about the experiment meant to probe whether they had read the instructions
 5 One question was which university the experiment was being run by, and the other
 6 question was how many trials there would be in the experiments. The 12 items along
 7 with 16 fillers were then randomly presented in a Latin square design after three
 8 practice trials. For each trial, participants read a short Mandarin paragraph which either
 9 established a context where a contrasting type is present (e.g., (3a), in which there are
 10 no dogs but there are people), or a context where the contrasting type is absent (e.g.,
 11 (3b), where there are neither people nor dogs in the park). Then participants listened to a
 12 sentence relevant to this setting with either a stressed noun or a stressed numeral-
 13 classifier combination. The task for the participants was to judge whether what they
 14 heard could fit the context that they read. They were asked to click either *consistent* or
 15 *inconsistent* based on their own judgements. After submitting the answer, a participant
 16 could move onto the next question. The whole survey was self-paced. It took less than
 17 30 minutes for the participants to finish the survey.

18 ***Predictions***

19 In the context where the type alternative is present (3a), we expected that the critical
 20 sentence with prosodic focus on the noun would be less consistent with the context,
 21 since prosodic focus on the noun (i.e., "I didn't even see one *dog*") should license the
 22 inference that the speaker didn't see anything else either, including the type alternative
 23 (i.e., "I didn't even see one *dog*, let alone one person"). Thus, we expected a difference
 24 in consistency ratings between the two prosodic conditions in this context. On the other
 25 hand, in the context where the type alternative is absent (3b), we expected no difference

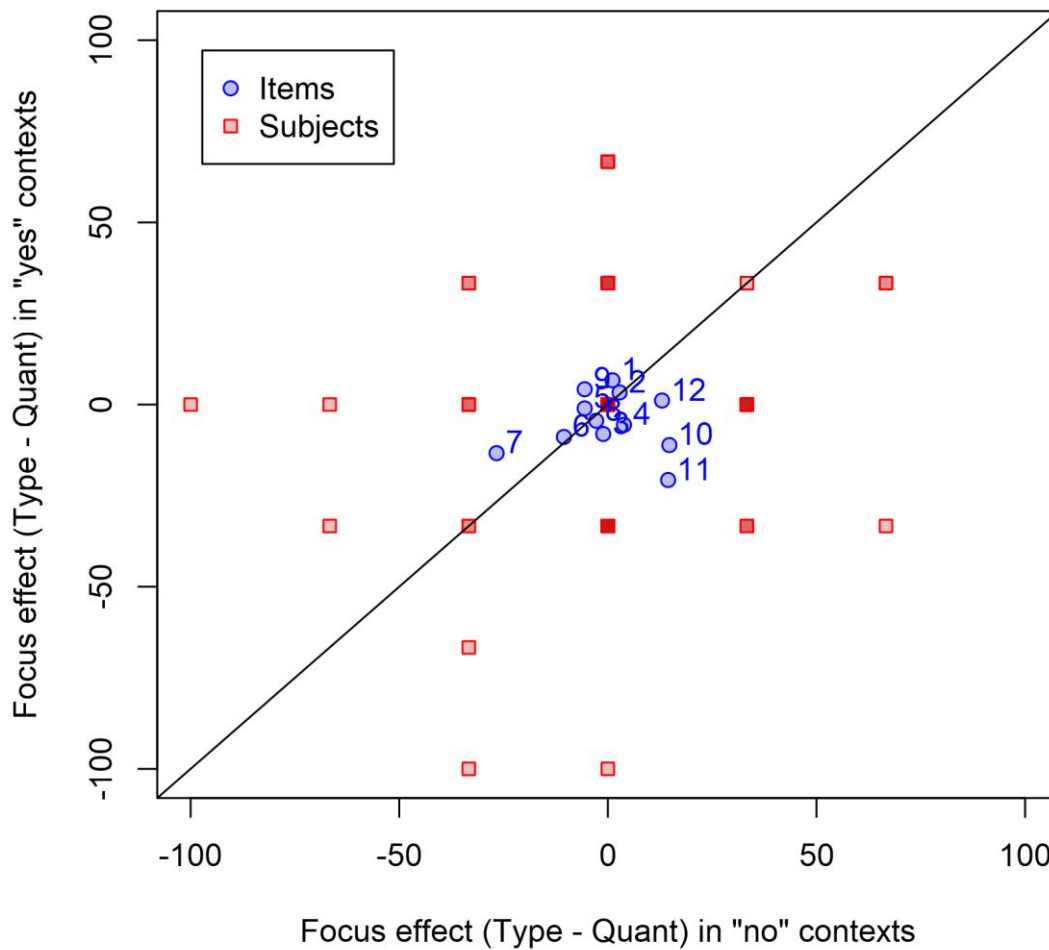
1 in consistency ratings between the two prosodic conditions, since both inferences (i.e.,
2 "I didn't even see one dog, let alone one person" and "I didn't even see one dog, let
3 alone two") are consistent with the context in which there are neither dogs nor people in
4 the park.

5

6 ***Results***

7 The full dataset and analysis code (for the R statistical programming environment) are
8 available at <https://osf.io/nsgfv/>.

9 Overall, in contexts where the alternative referent was present, participants
10 accepted 85.6% of sentences with quantifier stress and 82.1% of items with noun stress,
11 an difference in the expected direction; also consistent with the predictions, they
12 showed less difference in acceptance of different prosody in the context where the
13 alternative referent is also not available, accepting 89.6% of sentences with quantifier
14 stress and 90.0% of sentences with noun stress. Figure 1 shows the variability of the
15 effect across items (by-subject aggregates are not plotted; since each participant only
16 saw a small number of items and thus only had a small number of possible outcomes
17 per condition [0, 33, 66, or 100%], there is little variability to be seen).



1

2 *Figure 1. Effects of prosody in Experiment 1. Each point represents one stimulus item*
3 *or one participant. The x-axis represents the difference in percentage acceptance for*
4 *noun stress vs. quantifier stress prosody in contexts where the alternative referent is not*
5 *present (i.e., when neither dogs nor people were in the park), such that negative values*
6 *indicate when noun stress prosody was accepted less than quantifier stress prosody. The*
7 *y-axis shows this same difference, but in contexts where the alternative referent is*
8 *present (i.e., when there were no dogs in the park but there were people). Because the*
9 *prediction was that there would be a larger prosody effect (in the negative direction) in*
10 *these context than in contexts where the alternative is available, this means that points*

1 below the diagonal represent items showing effects consistent with the prediction, and
 2 points above the diagonal are inconsistent with the prediction. Note that at several
 3 places there are multiple subjects with points in the same location; these can be
 4 recognized by the darker square backgrounds (since the background coloring is
 5 opaque). For clarity, point labels are provided only for items, not for subjects.

6

7 The results were statistically analyzed using generalized (binomial) mixed-effects
 8 models with crossed random effects for subjects and items [21]. The predictors
 9 PROSODIC STRESS (*noun stress* vs. *quantifier stress*) and CONTEXT (alternative type
 10 present vs. alternative type absent) were sum-coded (as 0.5 and -0.5) and used as fixed
 11 predictors, along with their interaction; random effects of these three parameters were
 12 also fit for items [22], but not for subjects, since each subject had too few trials to fit
 13 this complex structure well. The significance of the crucial PROSODIC STRESS *CONTEXT
 14 interaction was assessed with a log-likelihood test comparing this model to a maximally
 15 similar model without the fixed interaction effect. The interaction did not reach
 16 significance in this comparison ($\chi^2(1)=0.14$, $p=.707$).

17

18 **Discussion**

19 The results of the experiment showed a numerical trend in the predicted direction, such
 20 that prosody influenced sentence acceptability in contexts where the alternative type
 21 was present and less so in contexts where the alternative type was absent. However, this
 22 trend was not statistically significant. Furthermore, even in contexts where prosody
 23 should have elicited an inference that does not fit the context (i.e., "I didn't even see one
 24 dog [let alone one person]", in a context where there were no people in the park),

1 sentence acceptance was still quite high, over 80%; this suggests that participants were
2 not influenced very much by prosody, as long as the lexico-semantic content of the
3 sentence fit the context. For this reason, we attempted to conceptually replicate the
4 experiment, while making changes to potentially increase the size of the effect. We
5 suspected that the binary nature of the acceptability judgment may have forced
6 participants to 'accept' sentences even when they were aware of slight inconsistencies;
7 thus, in this experiment we instead had participants rate sentences on a 6-point Likert
8 scale, which we predicted might allow them to register their awareness of the prosodic
9 mismatch and thus might increase the chances of observing a prosodic effect.
10 Otherwise, the predictions for Experiment 2 are the same as for Experiment 1: we
11 expect worse ratings for noun-stress prosody than for quantifier-stress prosody in
12 contexts where the alternative type is present, but not in contexts where the alternative
13 type is absent.

14

15 **Experiment 2**

16 *Participants*

17 Seventy-eight native speakers of Mandarin (63 users of traditional Chinese characters,
18 15 users of simplified Chinese characters) attended this experiment. Ten were excluded
19 for answering baseline questions incorrectly; the exclusion criteria and data collection
20 stopping rule were pre-registered at
21 <https://osf.io/bz6c2/register/5771ca429ad5a1020de2872e>. This left 68 participants (aged
22 18-70, mean 42) in the final analysis.

23 *Materials*

24 The materials are the same as those from Experiment 1.

1 *Procedure*

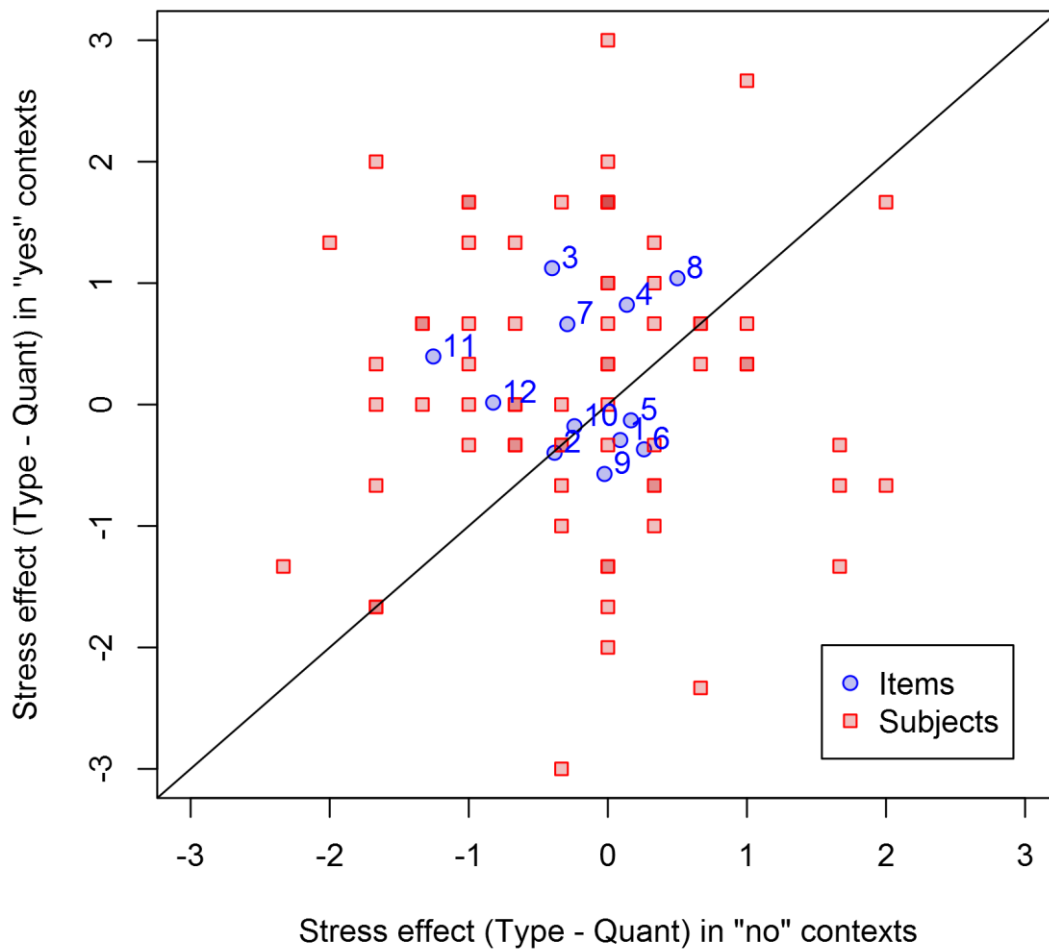
2 The procedure was the same as in Experiment 1, except that the task for the participants
 3 in Experiment 2 was to rate the consistency between what they heard and what read
 4 based on a 1-6 scale. 1 stood for *completely inconsistent*, while 6 stood for *completely*
 5 *consistent*. The participants were guided to go through two practice trials: one practice
 6 is an example of *completely inconsistent*, while the other is a practice of *completely*
 7 *consistent*, before starting the experiment. The example of *completely consistent* is
 8 provided in (6), where the audio content emphasized the quantity of water and has no
 9 conflicts with the written content.

- 10 (6) Written content: *Huang laoshi tongchang he henduo sui. Ta jintian hen*
 11 *mang. Ta mei he sui ye mei he kele. Ta de xuesheng gen ni shuo*: ‘Mr.
 12 Huang usually drank a lot of water. He was very busy today. He drank
 13 neither water nor coke. His student said:’
 14 Audio content: *Ta yi di sui dou mei he*. ‘He didn’t eat even **one drop of** water.’

15 *Results*

16 The full dataset and analysis code are available at <https://osf.io/nsgfv/>. In
 17 contexts where the alternative type is present, the mean consistency rating was 4.9 for
 18 sentences with noun-stress prosody and 4.7 for sentences with quantity-stress prosody;
 19 this difference is opposite the predicted direction. In contexts where the alternative type
 20 is absent, consistency ratings were 5.2 for noun-stress prosody and 5.4 for quantity-
 21 stress prosody. In the two contexts, the mean consistency rating was fairly high. The
 22 distribution of differences by subjects and items is shown in Figure 2. Since the effects
 23 were opposite the predicted direction, inferential statistics were not conducted.

24



1

2 *Figure 2. Effects of prosody in Experiment 2. Blue points represent items and red*
 3 *squares represent subjects; as in Experiment 1, points below the diagonal are subjects*
 4 *or items with differences in the predicted direction. Note that at several places there are*
 5 *multiple subjects with points in the same location; these can be recognized by the*
 6 *darker square backgrounds (since the background coloring is opaque). For clarity,*
 7 *point labels are provided only for items, not for subjects.*

8 **Discussion**

9 Experiment 2 failed to replicate the trend observed in Experiment 1. We suspected that
 10 the effects of prosody may have been weakened or obscured in these experiments by

1 two factors. First, the experimental contexts did not particularly draw participants'
 2 attention to prosody, and in fact may have drawn their attention more to lexico-semantic
 3 factors. Since the experiment included fillers in which the target sentence clearly
 4 mismatched the context based on basic semantics as in (7), many participants' attention
 5 may have been focused more on these issues. This type of fillers might become the
 6 standard for *completely inconsistent* for participants. Therefore, participants may have
 7 considered sentences with inconsistent prosody but consistent semantics to be fairly
 8 acceptable by comparison.

9 (7) Reading content: *Wangfang shi chuan le liang jian yifu, zuihou meiyou*
 10 *mai. Dianyuan gen ni shou:* 'Fang Wang tried on two pieces of clothes.
 11 She didn't buy any. The shop assistant told you.'

12 Audio content: *Ta yi jian yifu dou mei mai.* 'She didn't buy even one piece
 13 of clothes.'

14
 15 Secondly, Experiments 1 and 2 tested the effects of prosody on inferences indirectly, by
 16 testing whether prosody engenders an inference which mismatches a context (rather
 17 than by directly testing whether prosody engenders a given inference at all). In the two
 18 experiments, both prosody and lexico-semantic contents might influence the
 19 participants' judgements. Thus, the results are not merely reflective of prosody. In
 20 Experiment 3 we attempted to address these issues by using a more direct approach, and
 21 by using a design meant to explicitly draw participants' attention to prosody.

22 **Experiment 3**

23 ***Participants***

24 Sixty-four native speakers of Mandarin (63 users of traditional Chinese characters, 1

1 user of simplified Chinese characters) attended this experiment. Eleven were excluded
 2 for answering baseline questions incorrectly, and four for having low accuracy in the
 3 unambiguous filler trials. This left 49 participants (aged 18-60, mean 26) in the final
 4 analysis.

5 **Materials**

6 The experiment consists of 12 critical sentences along with 6 fillers. The critical
 7 sentences are in the format of (2). The participants were asked to listen to the same
 8 sentence in two kinds of prosodic patterns: one with stress on the noun, one with stress
 9 on the numeral-classifier combination. Afterward, they were asked to choose the most
 10 appropriate answer to be the first clause of a two-clause sentence. The question appears
 11 in the format as (8)a or (8)b. (8)a provides an alternative in the category of types, whereas
 12 (8)b offers an alternative in the domain of quantity.

13 (8) a. _____, gengbieshuo you guke le.
 14 much less there be customer ASP

15 ‘_____, much less customers.’ [an alternative in type]

16 b. _____, gengbieshuo you yi qun maomi le.

17 much less there be one CLF cat ASP

18 ‘_____, much less a group of cats.’ [an alternative in quantity]

19 This version of the experiment only had two conditions: follow-up contexts which stress
 20 the type alternative, and follow-up contexts which stress the quantity alternative. We
 21 predicted that sentences with stress on the noun (consistent with type focus) would be
 22 selected more often when the follow-up sentence stresses the type alternative (8) than

1 when it stresses the quantity alternative (8). The items were organized into two lists in a
 2 Latin square design. There were six fillers, which also appear in the same format.

3 Among the fillers, three of them were in positive polarity environments, and three
 4 of them are in negative polarity environments. For each trial, two audio files were
 5 provided: one option matches the follow-up context (9), while the other mismatches the
 6 follow-up context (9).

7 (9) _____, *genbei shuo xiao gongyu le*. ‘_____, much less a small
 8 apartment.’

9 Audio files:

10 a. *Ta mai de qi chengshi li de da haozai....* ‘He can afford a mansion in
 11 the city...’

12 b. *Ta chi de qi niupai...* ‘He can afford steaks...’

13 The fillers, which were unambiguous, also served to check the validity of the responses.

14 The full list of stimuli is available at <https://osf.io/nsgfv/>.

15 ***Procedure***

16 This experiment was administered via Ibex Farm. The 12 critical items along with 6 fillers
 17 were presented in a fully random order after two practice trials. The practices were
 18 designed to direct participants’ attention to prosodic differences. As in (10), the two audio
 19 files have the same format, but the placement of a contrastive stress determined the item
 20 to be contrasted. According to the written context provided in (10), only the prosody of
 21 (10) can match the follow-up sentence.

1 (10) Reading context: _____, *bu shi Xiaohan hui.* ‘_____, not
 2 Xiaohan who is able to.’

3 Audio files: a. *wo zhidao ta hui tiaowu* ‘I know it is she who is able to
 4 dance.’ [contrastive stress on *ta* ‘she’]
 5 b. *wo zhidao ta hui tiaowu* ‘I know it is cooking that she is able to do.’
 6 [contrastive stress on *hui tiaowu* ‘be able to dance’]
 7

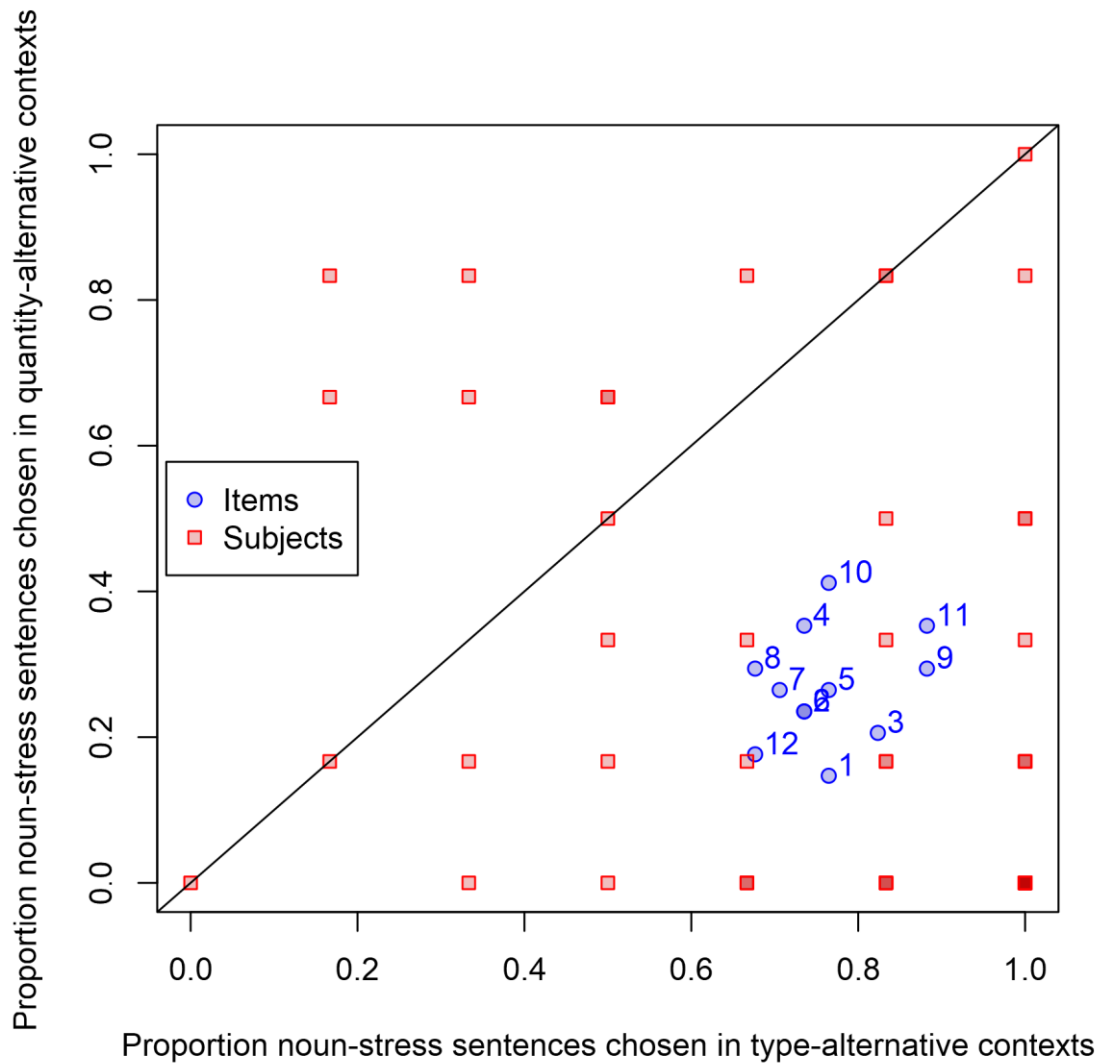
8 For each trial, with a written context sentence and two audio clips occurred on the
 9 screen at the same time. The task for the participants is to choose one of two audio clips
 10 to complete the sentence shown on the screen, which only the second clause of a two-
 11 clause sentence is provided. Participants could play the audio clips more than one time.
 12 The self-paced survey took less than 30 minutes to finish.

13

14 **Results**

15 The data and analysis code are available at <https://osf.io/nsgfv/>.

16 As shown in Figure 3, sentences with noun stress were chosen more often in
 17 contexts that evoked the type alternative than in contexts that evoked the quantity
 18 alternative; conversely, sentences with quantifier stress were chosen more often in
 19 contexts that evoked the quantity alternative than contexts that evoked the type
 20 alternative.



1

2 *Figure 3. Effects of context in Experiment 3. Each point represents, for a given stimulus*
 3 *item or participant, the proportion of trials in which the sentence with stress on the*
 4 *noun was selected. Since we predicted more selection of noun-stress sentences in*
 5 *contexts evoking the type alternative than in contexts evoking the quantity alternative,*
 6 *that means the prediction is that the mean should be below the diagonal line.*

7

8 The context effect was analyzed with a generalized (binomial) mixed-effects model
 9 regressing the binary response (coded with quantifier stress as the baseline level) on the
 10 fixed effect of context (dummy-coded with quantity-alternative contexts as the baseline

level) and maximal random effects for subjects and items. This model revealed a significant effect of context ($b=358$, $z=5.83$, $p<.001$), indicating that the likelihood of selecting a sentence with stress on the noun was significantly higher in type-alternative contexts than in quantity-alternative contexts.

General Discussion

This study tested the hypothesis about whether prosody is a valid determinant for pragmatic judgments. The connections among focus, prosody, and pragmatic inferences have been hinted in different literature [9-10, 15-18, 20], but the relations have not yet been specified. In order to prove the existence of the connection, we tested whether the type-contrast and quantity-contrast prosodic patterns can guide Mandarin native speakers to the correspondent scalar inferences. The results of the experiments show that Mandarin native speakers rely on prosody to build pragmatic reasoning, but they may not be aware of it. The sentences of stimuli appear in the same syntactic structure, which has the numeral ‘one’ and the classifier specified. This syntactic pattern inherently entails the semantics of quantity. According to the participants’ responses, they tend to use the quantity contrast for this syntactic structure if they are not aware of the prosodic changes. The effects of prosody were not strong in Experiment 1 and Experiment 2 because in this setting most of participants were not aware of prosodic variation and thus made their judgments based on syntax and semantics instead of prosody. In this case, the participants judged the consistency between what they heard and what they read based on the quantity-contrast inferences. The phenomenon shows that pragmatic inferences of a quantity contrast are defaulted in this syntactic construction. However, for those who were aware of the different patterns of prosody, their judgements of scalar inferences were based on the patterns of prosody. Their responses were consistent with our prediction. It shows that prosody is a crucial factor

1 for Mandarin native speakers to build their scalar reasoning, but the participants
 2 responded more to the syntax and semantics than to the prosody in the setting of
 3 Experiments 1 and 2. For the purpose of raising the awareness of the prosodic variation,
 4 the participants were guided to compare two prosodic patterns in Experiment 3.
 5 Moreover, the contextual information is minimalized in Experiment 3 in order to direct
 6 the participants' attention to the variation of prosody. The participants had to pick up
 7 one of the two audio files to complete the first clause of a two-clause sentence in a
 8 pragmatically coherent way. The design of the comparison successfully increased
 9 participants' awareness of the differences in prosody and guided the participants made
 10 judgments primarily based on prosody. As shown in the results of Experiment 3, the
 11 pragmatic inferences were made based on the item highlighted in the prosody. It shows
 12 that prosody is critical in constructing pragmatic inferences if other interfering elements
 13 are minimalized.

14 The experiments also provide the evidence for the observed connection between
 15 prosodic stresses and minimizers in the literature. The occurrence of a prosodic stress
 16 contributes to inducing a set of pragmatic inferences coherent with contexts. The
 17 placement of a prosodic stress is an indicator of where the attention of the native
 18 Mandarin speakers would be. This relation between prosodic prominence and loci of
 19 attention helps to account for the concept of focus in the syntax in Mandarin Chinese.

20

21 **Conclusion**

22 The present study showed that different prosodic patterns can guide hearers to
 23 induce different scalar reasoning. It has been observed that 'one'-phrases minimizers in
 24 numeral-classifier languages have two types of scalar inferences due to the structure as
 25 a numeral phrase [15, 16]. The two types of inferences, quantity-contrast and type-

1 contrast, are reflected in morphology in other numeral-classifier languages, but not in
 2 Mandarin. In the conditions where syntax, semantics, and morphology do not
 3 differentiate types of scalar inferences, native Mandarin speakers depend on prosody to
 4 determine the entailed conceptual scale. It has been noted that the numeral ‘one’ of
 5 Mandarin minimizers may bear a stress, but the actual loci of the stress and the purpose
 6 the stress were not specified. The experiments confirm that the loci of a prosodic stress
 7 carry pragmatic information and reveal the role of prosody in evoking pragmatic
 8 inferences.

9 Although the placement of a prosodic stress specifies the types of inferences, the
 10 induced scalar inferences are asymmetrical. The quantify-contrast inferences are easier
 11 for people to reach a consensus due to the involvement of a numeral phrase, while
 12 people might have different interpretations for the type-contrast inferences since they
 13 have different expectation of the minimal size and frequency of occurrence. How to
 14 determine the most accessible scalar interpretation will be the future extension of the
 15 current study.

16 References

- 17 [1] Morris, Charles. 1938. *Foundations of the Theory of Signs*. (*International*
 18 *Encyclopedia of Unified Science*, vol. 1, no. 2.) Chicago, IL: University of Chicago
 19 Press.
- 20 [2] Grice, Paul. 1989. *Studies in the Way of Words*. Cambridge: Harvard University
 21 Press.
- 22 [3] Horn, Laurence and Gregory Ward. 2005. *The Handbook of Pragmatics*. Blackwell.
- 23 [4] Fillmore, Charles J., Paul Kay, and Mary Catherine O’Connor. (1988). Regularity
 24 and idiomaticity in grammatical constructions: The case of let alone. *Language* 64:
 25 501- 38.
- 26 [5] Chao, Yuen Ren. 1933. Tone and intonation in Chinese. *Bulletin of the Institute of*
 27 *History and Philology, Academia Sinica* 4(3), 121-134.
 28
- 29 [6] Shih, Chilin. 1988. Tone and intonation in Mandarin. *Working Papers of the Cornell*
 30 *Phonetics Laboratory* 3, 83-109.

- 1 [7] Yuan, Jiahong. Intonation in Mandarin Chinese: Acoustics, perception, and
2 computational modeling. PhD diss., Cornell University, Aug., 2004.
3
- 4 [8] Giannakidou, Anastasia. 2011. Positive polarity items and negative polarity items:
5 variation, licensing, and compositionality. In *Semantics: An international handbook*
6 *of natural language meaning*, ed. by Claudia Maienborn, Klaus von Heusinger, and
7 Paul Portner, 1660-1712. Berlin: Mouton de Gruyter.
- 8 [9] Israel, Michael. 2011. *The Grammar of Polarity: Pragmatics, Sensitivity and the*
9 *Logic of Scales*. Cambridge: Cambridge University Press.
- 10 [10] Haspelmath, Martine. 1997. *Indefinite Pronouns*. Oxford/ New York: Oxford
11 University Press.
- 12 [11] Fillmore, Charles J., Paul Kay, and Mary C. O'Connor. 1988. Regularity and
13 idiomaticity in grammatical constructions: The case of *let alone*. *Language* 64 (3):
14 501-538.
- 15 [9] [12] Fauconnier, Gilles. 1975. Pragmatic scales and logical structure. *Linguistic*
16 *Inquiry* 6: 353-375.
- 17 [10] [13] Horn, Laurence R. 1989. *A Natural History of Negation*. Chicago: The
18 University of Chicago Press.
- 19 [11] [14] Ahrens, Kathleen, and Chu-Ren Huang. 2016. Classifiers. In *A reference*
20 *grammar of Chinese*, ed. Chu-Ren Huang and Dingxu Shi. Cambridge: Cambridge
21 University Press.
- 22 [15] Lee, Chungmin. 2006. Contrastive topic/focus and polarity in discourse. In *Where*
23 *Semantics Meets Pragmatics*, ed. by Klaus von Heusinger and Ken Turner, 381-420.
24 Oxford: Elsevier Science.
- 25 [16] Nakanishi, Kimiko. 2006. The semantics of *even* and negative polarity items in
26 Japanese. In *The Proceedings of the 25th West Coast Conference in Formal*
27 *Linguistics (WCCFL 25)*, ed. by Donald Baumer, David Montero and Michael
28 Scanlon 288-296. Somerville, MA: Cascadia Press.
- 29 [17] Zhang, Ning. 2000. Object shift in Mandarin Chinese. *Journal of Chinese*
30 *Linguistics* 28 (2), 201-246.
- 31 [15] [18] Tsai, Wei.-Tien Dylan. 2004. On formal semantics of *zhi* and *lian* in Chinese.
32 *Zhouguo Yuwen* 中國語文 [Chinese Language] 2, 99-111.
- 33 [19] Huang, C-T, James, Y.-H. Audrey Li and Yafei Li. 2009. *Syntax of Chinese*.
34 Cambridge: Cambridge University Press.
- 35 [20] Chao, Yuen-Ren. 1968. *A Grammar of Spoken Chinese*. Berkeley: University of
36 California Press.
- 37 [21] Baayen, H., Davidson, D., & Bates, D. (2008). Mixed-effects modeling with
38 crossed random effects for subjects and items. *Journal of Memory and Language*, 59,
39 390-412.

- 1 [22] Barr, D., Levy, R., Scheepers, C., & Tily, H. (2013). Random effects structure for
- 2 confirmatory hypothesis testing: keep it maximal. *Journal of Memory and Language*,
- 3 68, 255-278.