

The Epistemic Indefinite *shenme* in Mandarin

Zhuo Chen

1. Introduction

Wh-phrases such as *shei* ‘who’ and *shenme* ‘what’ in Mandarin Chinese have both interrogative uses and non-interrogative uses. When used non-interrogatively, they are licensed in NPI contexts when interpreted as *somebody* and *something*, as is shown in (1a) and (1b), which earned them the name of ‘existential polarity *wh*-items’. However, they are also licensed in certain environments that a typical NPI like English *any* is allergic to. Two such examples are given in (1c)¹ and (1d). In order to unify the licensing environments, Lin (1998) proposed that a NEEC (non-entailment of existence condition) is imposed by the polarity *wh*-items on their contexts. This condition requires that the proposition in which an existential polarity item appears does not entail the existence of a referent satisfying the description of the existential polarity item². Xie (2007) examined a wide range of the distribution of the polarity *wh*-items and concluded that they are licensed in non-veridical environments (Giannakidou, 2002).

- (1) a. Wo mei mai shenme (dongxi)
I not buy what (thing)
‘I didn’t buy anything.’
- b. Ruguo shei qifu ni, jiu gaosu wo
If somebody bully you, then tell me
‘If somebody bullies you, tell me.’
- c. Keneng shei qifu ta le
Possibly who bully him Asp
‘Possibly, somebody bullied him.’
*‘Possibly, anybody bullied him.’
- d. Zhangsan yiwei wo mai le shenme
Zhangsan think I buy Asp what
‘Zhangsan thought I bought something.’
*‘Zhangsan thought I bought anything.’

A prediction from Lin (1998) and Xie’s (2007) account is that the polarity *wh*-phrases are intolerable in simple affirmative episodic sentences. However, we find sentences like (2) in Mandarin.

- (2) Zhangsan mai le san ben shenme shu
Zhangsan buy Asp three CL what book
‘Zhangsan bought three books of a certain kind, (but I don’t know what kind it is).’

If *shenme* as an existential polarity item obeys NEEC, (2) should be ungrammatical since there exist referents satisfying the description of the polarity item. The non-veridicality account will also exclude (2) because positive episodic sentence is a veridical context. It is thus surprising to find the occurrence of *shenme* in (2) if we follow the extant analyses of the polarity *wh*-items in Mandarin. In terms of the interpretation, *shenme* as is used in (2) gives rise to an epistemic ignorance effect on the part of the speaker. Specifically, it signals ignorance of a kind, as is shown in the gloss in (2).

* Zhuo Chen, The Graduate Center, CUNY, zchen3@gradcenter.cuny.edu. I want to thank Sam Alxatib for discussion.

¹ Abbreviations in the glosses of Chinese examples: Asp: aspect marker; CL: classifier.

² Lin (1998) later refined this condition in order to account for the locality constraints.

The remainder of the paper is organized as follows. Section 2 introduces the kind reading of *shenme*. Section 3 discusses the epistemic indefinites from other languages and reviews Weir's (2012) proposal on the epistemic use of *some* in English. Section 4 gives a semantic account for the epistemic indefinite *shenme* and section 5 concludes the paper.

2. The Kind Reading of *Shenme*

The idea that *shenme* has a kind reading apart from its individual reading is not exotic. First, Heim (1987) presented some evidence from the definiteness effect in *wh*-questions that *what* — the English counterpart of *shenme* — introduces a kind variable and is interpreted as *something of what kind*. Second, Lin (1999) studied the phenomenon of double quantification in Chinese bare conditionals and argued that *shenme* introduces a kind variable, lending support to Heim's hypothesis.

2.1. Heim (1987)

A well-known fact in English is that strong quantifiers are not licensed in there-insertion sentences. Hence the ungrammaticality of (3).

- (3) *There is every student.

Given the fact that the there-insertion sentences do not tolerate overt variable pronouns as shown below in (4), Heim (1987) hypothesized that overt variable pronouns are like strong NPs.

- (4) *Few people_{*i*} said there would be them_{*i*} at the party.

If other variables behave in the same way, we will predict that the variables left behind by *wh*-movement are not allowed in there-be sentences. However, within *wh*-questions, we seem unable to find a consistent pattern with regard to this prediction. Heim noticed an interesting contrast between (5) and (6). While we find the definiteness effect in (5), we don't find it in (6).

- (5) Which actors were there ??in the room?

- (6) What is there in Austin?

In order to explain the felicity of (6), Heim proposed that *wh*-movement in *what* questions is in fact a kind of pied-piping in disguise. *What* should be interpreted as *something of what kind*. In a *wh*-question, the interrogative operator binds the kind variable introduced by *what* but the whole *wh*-phrase *something of x kind* is reconstructed back to its original position for interpretation. Since *something of x kind* is an indefinite, no definiteness effect would emerge.

The kind reading of *what* is supported by Lin's (1999) analysis of the bare conditionals in Chinese, to which now we turn.

2.2. Lin (1999)

In a Chinese bare conditional like (7), the correct interpretation is "if you break something of a certain kind, you need to go and buy something of the same kind to compensate". To make things simple, the intensional semantics are ignored here and a crude representation is given in (8). The superscripts specify whether the variable is a kind-level variable or an individual-level variable. The realization relation *R* holds between a kind and an individual instance of the kind.

- (7) Ni dapo shenme, jiu dei qu mai shenme lai pei
 you break what, then must go buy what to compensate
 'If you break something of kind *x*, then you must go to buy something of kind *x* for compensation.'

- (8) $\forall x^k [\exists y^s [R(y^s, x^k) \ \& \ \textit{break}(\textit{you}, y^s)]] [\exists z^s [R(z^s, x^k) \ \& \ \textit{go-to-buy}(\textit{you}, z^s)]]$

What *shenme* does in the bare conditional, as argued by Lin, is to introduce a kind-level variable. Since there is no other adverbial quantifiers present in the sentence, following Heim (1982), Lin assumes that a covert universal quantifier binds the kind-level variable introduced by *shenme*. When the stage-level predicates like *dapo* and *mai* applies to a kind, they introduce existential quantification over an individual-level variable that stands in a realization relation to that kind. This is how the intuitive existential force of *shenme* in the sentence is captured and the intended interpretation for (7) is obtained.

Thus far, we have seen that *shenme* has a kind reading. The lack of definiteness effect in *what* questions and the double quantification in Chinese bare conditionals both speak to this fact. The remaining task for us in order to explain the reading in (2) is to derive the epistemic ignorance effect of *shenme*. Before we give a semantic account of *shenme*, let's take a look at the epistemic indefinites from other languages that signal the speaker's ignorance, just like what *shenme* does in (2).

3. Epistemic indefinites

3.1. Epistemic indefinites cross-linguistically

Cross-linguistically, we find epistemic indefinites that signal the speaker's ignorance (or indifference) as to the referent of the NP containing the epistemic indefinite. Examples include Spanish *algún*, German *irgendein* and English *some*. The German sentence (9) and Spanish sentence (10) have the same meaning as the English sentence (11). The use of epistemic indefinites in these sentences prevent the speaker from identifying the doctor that Maria married.

(9) Maria hat irgendeinen Arzt geheiratet (# und zwar Dr. Smith)

(10) María se casó con algún médico (# en concreto con el Dr. Smith)

(11) Maria married some doctor or other, (# namely Dr. Smith)

Alonso-Ovalle & Menéndez-Benito (2010) came up with the following lexical entry for *algún* to account for the epistemic effect generated by it.

(12) $[[\textit{algún}]] = \lambda f_{\langle et, et \rangle} \lambda P_{\langle e, t \rangle} \lambda Q_{\langle e, t \rangle} : \textit{anti-singleton}(f). \exists x [f(P)(x) \ \& \ Q(x)]$

Built in the semantic meaning of *algún* is an anti-singleton subset selection function f . When f takes a set as its argument, it returns a subset of that set. If the subset has at least two members in it, the anti-singleton presupposition of f is satisfied and we can proceed with the semantic composition. If the subset is a singleton set, the presupposition fails and the composition crashes. In (10), the sentence claims María married a doctor in the subset that f picks out from the set of all doctors and *algún* signals that the subset is not a singleton set. What the anti-singleton presupposition does is to add this epistemic effect that the speaker has in mind more than one doctors and s/he knows Maria married one of them but s/he does not know exactly who.

The anti-singleton presupposition is demonstrated in the incompatibility between *algún* and superlatives. Alonso-Ovalle and Menéndez-Benito gave sentence (13) as an example. (13) is doomed since the superlative selects only one book, directly contradicting the anti-singleton presupposition.

(13) #Juan compró algún libro que resultó ser el más caro de la librería
 Juan bought ALGÚN book that happened to be the most expensive in the bookstore
 'Juan bought a book that happened to be the most expensive one in the store.'

3.2. English *some*

English *some* performs similarly to Spanish *algún* when it is combined with a singular NP. It also signals the speaker's ignorance as to the witness of the existential phrase. Hence the infelicity of the continuation in the parenthesis in (14).

(14) Some tomato in this basket went bad. (#Namely, that one).

Given this similarity, Weir (2012) extended Alonso-Ovalle and Menéndez-Benito's analysis to English *some*. Besides the ignorance of the individual witness, however, Weir found that *some* can also denote speaker's ignorance of a subkind. An example of the such reading is shown in (15).

(15) I came home to find some plant growing through a hole in my wall.

Sentence (15) cannot mean 'I came home to find a plant growing through a hole in my wall, but I don't know which plant it is'. The speaker eyeballed the plant and thus definitely knew which plant it was. The intuitive interpretation of this sentence is more like 'I came home to find a plant growing through a hole in my wall, but I don't know what kind of plant it is'. In other words, the speaker cannot decide what subkind of plant this plant belongs to.

To account for this epistemic effect of subkind ignorance, Weir (2012) followed Krifka (1995) in assuming that English has two covert classifiers. One is the individual classifier CL_{ind} and the other is the kind classifier CL_{kind} . A common noun root like \sqrt{plant} denotes the sum of all plants. When combined with a common noun, CL_{ind} will return a set of individuals realizing the kind denoted by the common noun and CL_{kind} will return a set of subkinds, all of which belong to the kind denoted by the common noun. The semantic denotation of the common noun and the two classifiers are given below, following Weir (2012).

$$(16) \llbracket \sqrt{plant} \rrbracket = \sum Plant$$

$$(17) \llbracket CL_{ind} \rrbracket = \lambda x \lambda y. kind(x) \ \& \ individual(y) \ \& \ y \Pi x$$

$$(18) \llbracket CL_{kind} \rrbracket = \lambda x \lambda y. kind(x) \ \& \ kind(y) \ \& \ y \Pi x$$

Common nouns in English are now ambiguous between an individual reading and a kind reading depending on which classifier it is combined with. *Some*, when it takes a common noun that is combined with an individual classifier, it generates the epistemic effect of individual ignorance as is shown above in (14). When it takes a common noun that is combined with a kind classifier, it generates the epistemic effect of subkind ignorance exemplified in (15). The lexical entry for *some* proposed by Weir is given below.

$$(19) \llbracket some \rrbracket = \lambda f_{\langle et, et \rangle} \lambda P_{\langle e, t \rangle} \lambda Q_{\langle e, t \rangle} : anti-singleton(f). \exists x[(f(P))(x) \ \& \ \exists y[y \Pi x \ \& \ Q(y)]]$$

Based on this entry for *some*, we can give the assertion and presupposition made by the sentence 'some plant grew through my wall' in (20)³.

$$(20) \llbracket some \ plant \ grew \ through \ my \ wall \rrbracket$$

Assertion: $\exists x[(f(\lambda y. kind(y) \ \& \ y \Pi (\sum Plant)))(x) \ \& \ \exists z[z \Pi x \ \& \ grew-through-my-wall(z)]]$

Presupposition: $|f(\lambda y. kind(y) \ \& \ y \Pi (\sum Plant))| > 1$

This sentence says that 'there is a subkind x of plants, and there is an individual plant z from the subkind x that grew through my wall and the speaker conveys that the set of subkinds from which x is selected from is not a singleton set.'

4. The semantic account of *shenme*

4.1. Differences between *shenme* and *some*

We are almost ready to account for the subkind ignorance reading in (2). Before we set out to do the task, we need to answer one question. When *shenme* signals the speaker's ignorance of a subkind, is it just a counterpart of English *some*? If this is the case, we should be able to use the lexical entry of *some* for *shenme*. However, the answer to this question seems to be negative.

³ The individual selection built in the lexical entry of *some* does not necessarily select an atomic individual z . Weir (2012) pointed out that the unknown-kind reading of *some* is number neutral.

Mandarin Chinese is a classifier language. Classifiers are used to individuate the kind denoted by the common noun (Krifka, 1995). Individual classifiers individuate a kind into its individual instances and kind classifiers individuate a kind into its subkinds. If *some* takes a common noun which is combined with a covert kind classifier in the subkind ignorance reading as is shown above in (15), with the assumption that *shenme* is just a counterpart of *some*, we will predict that *shenme* should take an overt kind classifier in Chinese to derive the same subkind ignorance reading. This direct transfer does not work, unfortunately. In English, a covert kind classifier is first combined with a common noun and returns a set of subkinds of the kind denoted by the common noun. *Some* then further selects individual instances from a subkind. In Mandarin, however, we find that *shenme* cannot be separated from the common noun by a classifier. This is why (21) is plainly bad.

- (21) *Zhangsan mai le yi shenme zhong shu
 *Zhangsan buy Asp one what CL_{kind} book

Moreover, *shenme* seems to be incompatible with kind classifiers. In (22), the intended interpretation is ‘Zhangsan bought some books which belong to three subkinds of a certain kind of book, but I don’t know what kind of it is’. The sentence, however, is not felicitous. This incompatibility does not surface when *shenme* takes an individual classifier, as is shown above in (2).

- (22) *Zhangsan mai le san zhong shenme shu
 Zhangsan buy Asp san CL_{kind} what book
 Zhangsan bought three subkinds of a certain kind of book, but I don’t know what kind it is.

Given the complementary distribution of *shenme* and kind classifiers, a reasonable hypothesis would be that *shenme* is a kind classifier itself and competes with other kind classifiers for one position in the syntactic structure. Different from the covert kind classifiers in English, however, *shenme* needs to introduce an epistemic ignorance effect. Following Alonso-Ovalle & Menéndez-Benito (2010), I attribute this epistemic effect to an anti-singleton requirement. *Shenme* not only individuates the kind denoted by the common noun into its subkinds, but also requires that the set of subkinds contain at least two members. This constraint a function imposes on its output is called a post-supposition⁴. *Shenme* introduces an anti-singleton post-supposition.

4.2. The semantics of *shenme*

From the differences between *shenme* and *some*, we can conclude that *shenme* has two roles. First, it is a kind classifier. When combined with a common noun, it individuates the kind denoted by the common noun into its subkinds. Second, it imposes an anti-singleton requirement on the output set of subkinds. When *shenme* is applied to a common noun, the output set of subkinds cannot be a singleton set. This anti-singleton requirement can be regarded as a post-supposition carried by *shenme*.

The proposed lexical entry for *shenme* is given in (23).

- (23) $\llbracket shenme \rrbracket = \lambda x \lambda y. kind(x) \ \& \ kind(y) \ \& \ y \Pi x$
 Post-supposition: $|\lambda y. kind(x) \ \& \ kind(y) \ \& \ y \Pi x| > 1$

This lexical entry successfully realizes the two roles of *shenme*. It takes a kind as an argument and returns a set of subkinds of that kind. With the post-supposition, it requires that there are at least two subkinds in the output set.

I assume that kind selection always precedes individual selection. An individual classifier always selects a kind before it further selects a plurality of individuals (or a single individual) whose size is determined by the numeral. The lexical entry for an example individual classifier *ben* (the classifier for ‘book’) is presented below.

⁴ Farkas (2002) constraint-based semantics for indefinites is the predecessor of the postsuppositional approach.

$$(24) \llbracket ben \rrbracket = \lambda n \lambda P_{\langle e,t \rangle} \lambda Q_{\langle e,t \rangle} . \exists x [P(x) \ \& \ \exists z (z \Pi x \ \& \ |z| = n) \ \& \ Q(z)]$$

For *san ben shenme shu* ‘three what book’, we have (25).

$$(25) \llbracket san \ ben \ shenme \ shu \rrbracket = \lambda Q . \exists x [\lambda y . kind(y) \ \& \ y \Pi (\sum book)](x) \ \& \ \exists z (z \Pi x \ \& \ |z| = 3) \ \& \ Q(z)]$$

Now we are ready to compose the meaning for sentence (2).

$$(26) \llbracket Zhangsan \ mai \ le \ san \ ben \ shenme \ shu \rrbracket = \exists x [\lambda y . kind(y) \ \& \ y \Pi (\sum book)](x) \ \& \ \exists z (z \Pi x \ \& \ |z| = 3) \ \& \ bought(z)(Zhangsan)]$$

To put it in words, the sentence means ‘there is a subkind x of books, and there are three books z that is part of the subkind x , and Zhangsan bought z , and the speaker wants to convey to the hearer that the set of subkinds of books that x is selected from is not a singleton set’. This is exactly the interpretation we get from (2): Zhangsan bought three books of a certain kind, but the speaker does not know what kind it is.

5. Conclusion

The previous accounts of the non-interrogative *wh*-phrases in Mandarin Chinese fail to capture their epistemic uses. In this paper, we have studied the epistemic use of a particular *wh*-item *shenme*. From its counterpart *what* in English and its use in the Chinese bare conditionals, we establish that *shenme* has a kind reading. Furthermore, based on the observation that *shenme* cannot be separated from the common noun and *shenme* is in complementary distribution with the overt kind classifiers in Chinese, we assume that *shenme* is a kind classifier itself that individuates the kind denoted by the common noun into its subkinds. In order to obtain the epistemic ignorance effect, we propose that *shenme* imposes a post-supposition on the output set of subkinds, that is, the output set should contain at least two members.

Further research should be done to lay out the distribution of the epistemic uses of the *wh*-phrases and a theory that can explain both the polarity uses and the epistemic uses is called for. The environments that license the polarity uses are relatively well-studied now. A question to ask is whether the complement of the polarity environments are exactly the environments which give rise to the epistemic effect. The answer seems to be no. Suppose John is talking about Mary, Sue and Beatriz. He knows that Mary bought a book, Sue bought a T-shirt and Beatriz bought a pen. In this context, it is appropriate for John to utter (27) even if he knows who bought what. *Shenme* does not convey an ignorance effect in this case. Alonso-Ovalle & Shimoyama (2014) observed the same cancellation of ignorance effect for Japanese *wh*-ka when it occurs in the upward-entailing nuclear scope of a universal quantifier. What kind of environments will trigger the ignorance effect of the non-interrogative *wh*-phrases, therefore, needs more research.

- (27) Mei ge ren dou mai le dian shenme
Every CL person DOU buy Asp CL what
‘Everyone bought something.’

Within the *wh*-words, we also find variation in their epistemic uses. For example, when *shei* ‘who’ is used to signal the speaker’s ignorance, we seem to always need a demonstrative ‘that’ before it. When the demonstrative determiner is absent, we only get the interrogative reading. When the demonstrative determiner is present, the use of *shei* signals the speaker’s ignorance as to the witness of the existential claim.

- (28) a. Zhangsan pengjian le shei
Zhangsan come across Asp who
‘Who did Zhangsan come across?’
b. Zhangsan pengjian le na ge shei
Zhangsan come across Asp that CL who
‘Zhangsan came across a person, (I forgot who it is).’

These remaining questions are left for future research. To conclude, the epistemic use of *wh*-items in Mandarin Chinese reveals that the previous accounts only paint a partial picture of their semantic properties. An investigation of the epistemic uses will give us a more comprehensive distribution of the *wh*-phrases in Chinese and help us find a common core in their semantics.

References

- Alonso-Ovalle, Luis & Paula Menéndez-Benito (2010). Modal indefinites. *Natural Language Semantics* 18:1, 1–31.
- Alonso-Ovalle, Luis & Junko Shimoyama (2014). Expressing ignorance in the nominal domain: Japanese *wh-ka*. Robert, Santana-Labarge (ed.), *Proceedings of West Coast Conference on Formal Linguistics*, Cascadilla Press, Somerville, MA, vol. 31, 11–20.
- Farkas, Donka (2002). Varieties of indefinites. Jackson, B. (ed.), *Proceedings of Semantics and Linguistic Theory*, vol. 12, 59–83.
- Giannakidou, Anastasia (2002). Licensing and sensitivity in polarity items: from downward entailment to nonveridicality. *Proceedings from the Annual Meeting of the Chicago Linguistic Society*, vol. 38, 29–53.
- Heim, Irene (1982). *The Semantics of Definite and Indefinite NPs*. Ph.D. thesis, University of Massachusetts, Amherst.
- Heim, Irene (1987). Where does the definiteness restriction apply? Evidence from the definiteness of variables. Reuland, Eric & ter Meulen, Alice G. B. (eds.), *The Representation of (In)definiteness*, Cambridge, MA: MIT Press, vol. 14, 21–42.
- Krifka, Manfred (1995). Common nouns: A contrastive analysis of Chinese and English. Carlson, G. & Pelletier, J. (eds.), *The Generic Book*, University of Chicago Press, 398–411.
- Lin, Jo-Wang (1998). On existential polarity *wh*-phrases in Chinese. *Journal of East Asian Linguistics* 7:3, 219–255.
- Lin, Jo-Wang (1999). Double quantification and the meaning of *shenme* what in Chinese bare conditionals. *Linguistics and Philosophy* 22:6, 573–593.
- Weir, Andrew (2012). Some, speaker knowledge, and subkinds. Rendsvig, R. K. & Katrenko, S. (eds.), *Proceedings of the ESSLLI 2012 student session*, 180–190.
- Xie, Zhiguo (2007). Nonveridicality and existential polarity *wh*-phrases in Mandarin. *Proceedings from the Annual Meeting of the Chicago Linguistic Society*, vol. 43, 121–135.

Proceedings of the 34th West Coast Conference on Formal Linguistics

edited by Aaron Kaplan, Abby Kaplan,
Miranda K. McCarvel, and Edward J. Rubin

Cascadilla Proceedings Project Somerville, MA 2017

Copyright information

Proceedings of the 34th West Coast Conference on Formal Linguistics
© 2017 Cascadilla Proceedings Project, Somerville, MA. All rights reserved

ISBN 978-1-57473-471-3 library binding

A copyright notice for each paper is located at the bottom of the first page of the paper.
Reprints for course packs can be authorized by Cascadilla Proceedings Project.

Ordering information

Orders for the library binding edition are handled by Cascadilla Press.
To place an order, go to www.lingref.com or contact:

Cascadilla Press, P.O. Box 440355, Somerville, MA 02144, USA
phone: 1-617-776-2370, fax: 1-617-776-2271, sales@cascadilla.com

Web access and citation information

This entire proceedings can also be viewed on the web at www.lingref.com. Each paper has a unique document # which can be added to citations to facilitate access. The document # should not replace the full citation.

This paper can be cited as:

Chen, Zhuo. 2017. The Epistemic Indefinite *shenme* in Mandarin. In *Proceedings of the 34th West Coast Conference on Formal Linguistics*, ed. Aaron Kaplan et al., 115-121. Somerville, MA: Cascadilla Proceedings Project. www.lingref.com, document #3303.