Specific unknowns: a case study of epistemic indefinites in Cantonese

Tommy Tsz-Ming Lee
University of Southern California
tszmingl@usc.edu

Jan 7-10, 2021 | LSA 2021 Virtual Annual Meeting

Contents

1 Introduction **2**

2 Epistemic indefinites in Cantonese: *mzi* + WH **3**
2.1 Cancellability, reinforcement, and the anti-singleton constraint 4
2.2 Obligatory wide intensional scope 4
2.3 Variable quantificational scope 5

3 Analysis **6**
3.1 Motivations for a conventional implicature approach 6
3.2 Motivations for a choice-functional analysis 7
3.3 Implementation 7

4 Evidence for grammaticalization **8**
4.1 Three positions of *m + zi* 8
4.2 Corpus data 9
4.3 Fusion of predicate and *wh*-expressions 10

5 Conclusions **10**

1 Introduction

This talk concerns how languages bundle an existential claim and an ignorance inference in a nominal expression. For example, the article *algún* in Spanish convey both meanings:

(1) #María se casó con **algún** estudiante del departamento de lingüística: en concreto con Pedro
María SE married with ALGÚN student of.the department of linguistics: namely with Petro
'María married a linguistics student, namely Pedro.' (Alonso-Ovalle and Menéndez-Benito 2010, p.2)

... as opposed to English *a/some*:

(2) Mary married **a/some** linguistics student, namely, Peter.

Indefinites that come with the speaker's ignorance over the witness of the indefinite are often regarded as **epistemic indefinites** (EIs, Alonso-Ovalle and Menéndez-Benito (2015)).

Across languages, the ignorance component is encoded via different morpho-syntatic means.

- article/determiner + NP (e.g. Spanish *algún* NP and German *irgendein* NP, i.a.)
- WH + particle (e.g. Japanese WH-*ka*, Tiwa WH-*khi*, i.a.)

The ignorance component display a non-uniform properties. Various characterizations have been suggested for the ignorance component (cf. Alonso-Ovalle and Menéndez-Benito 2013).

❶ Ignorance as a conversational implicature
❷ Ignorance as a result of a felicity condition or as a presupposition (non-Gricean approaches)

Different characterizations to EIs

❶ Ignorance as a conversational implicature

(a) Ignorance as a quantity implicature
A marker that imposes an anti-singleton constraint on the domain of the nominal
e.g. Spanish *algún* (Alonso-Ovalle and Menéndez-Benito 2010), Japanese *-ka* (Alonso-Ovalle and Shimoyama 2014)

(b) Ignorance as a manner implicature (*cf.* lexical blocking, McCawley 1978)
A marker that is in lexical competition with another expression
e.g. Tiwa *-khi* (Dawson 2018), Russian *-to* (Geist 2008)

❷ Ignorance as a result of a felicity condition or as a presupposition (non-Gricean approaches)

(c) Ignorance as a felicitous shift in identification methods
A marker that trigger an obligatory shift in identification method
e.g. German *irgendein*, Italian *un qualche* (Aloni and Port 2015), Czech *-si* (Šimík 2014), Sinhala *hari/də* (Slade 2015)

(d) Ignorance as intended referential vagueness
A marker that encodes anti-specificity
e.g. French *un quelconque* (Jayez and Tovena 2006), Greek *-dhipote* (Giannakidou and Quer 2013)

This talk is a case study on EIs in Cantonese, which represent a novel type of EIs in terms of ...

- the morpohological structure: *m* + *zi* + WH , literally, 'not + know + WH';
- the properties of the ignorancce component, which ...
  - is not cancellable or reinforceable;
  - survives intensional operators (i.e. denoting "specific unkonwns")
  - can scope below quantifiers and be distributed.

Claims on the ignorance component in Cantonese EIs:

- The nature : is **a conventional complicature** (Grice 1975; Potts 2005; Horn 2007), representing a third type of the ignorance component (different from ❶ and ❷);
- The source : it originates from the lexical meaning of *m-zi* 'not-know' and becomes a non-at-issue content due to **grammaticalization** of *m-zi* into **a choice function**.

➜ a less mentioned but important link between the nature and the source

▮ Road map: §2 Ignorance; §3 Analysis; §4 Grammaticalization; §5 Conclusion

2 Epistemic indefinites in Cantonese: *mzi* + WH

In Cantonese, EIs take the form of *mzi* + WH (henceforth, ***mzi*-indefinites**).

(3) Aaming tai-zo [**mzi** bin-bun syu], (# zikhai Hunglaumung)
Aaming read-PERF MZI which-CL book namely Dream.of.the.red.chamber
'Aaming read some book, namely, *Dream of the Red Chamber*. '

Morphologically, *mzi* is a combination of the negation *m* and the attitude verb *zi* 'know'.

(4) ngo **m-zi** [Aaming tai-zo bin-bun syu]
I not-know Aaming read-PERF which-CL book
'I don't know which book Aaming read.'

Mzi in (3) occupies a position unavailable to other predicates and is used as an adnominal modifier.

2.1 Cancellability, reinforcement, and the anti-singleton constraint

The ignorance component of *mzi*-indefinites do *not* show the signature properties of a conversational implicature.

(5) Non-cancellability (hence not calculable)

#Aaming tai-zo **mzi** bin-bun syu, ji ngo zidou hai bin-bun
Aaming read-PERF MZI which-CL book, and I know be which-CL
'Aaming read some book, and I know which (book it is).'

(6) Redundancy of Conjunction (Horn 1972); Non-reinforceability (Sadock 1978)

#Aaming tai-zo **mzi** bin-bun syu, ji ngo m-zidou hai bin-bun
Aaming read-PERF MZI which-CL book, and I not-know be which-CL
'Aaming read some book, and I don't know which (book it is).'

Additionally, one key property of deriving the ignorance via a quantity implicature concerns the anti-singleton constraint (Alonso-Ovalle and Menéndez-Benito 2010).

- It requires the domain of quantification to be non-singleton, which thus implicates the speaker's ignorance over the witness.

(7) Absence of the anti-singleton requirement

taihaa! **mzi** bin-go gaausau hai toi soengmin tiumou
look MZI which-CL professor at table top dance
'(Pointing at the professor) Look! Some professor is dancing on the table!'

✔ ostention; ✘ naming/descprtion

Note that *mzi* is compatible with ignorance concerning ostension, naming or description, i.e. the ignorance component is relatively unconstrained in this regard; see Appendix A.

2.2 Obligatory wide intensional scope

Another property of a *mzi*-indefinite concerns its interpretation with regard to intensional operators.

- The ignorance component is retained when embedded (i.e. it is *projective*).
- Notably, it is unambiguously scopally specific.

(8) Wide scope over attitude verbs

Aafan soeng tong **mzi** bin-go jisang jitfan
Aafan want with MZI which-CL doctor marry
'Aafan wants to marry to some doctor ... '
(i) ✔ ... they know each other for two years. scopally specific
(ii) ✘ ... but she does not know any doctor. scopally non-specific

The same applies to deontic modals as well:

(9) Wide scope over deontic modals

Aafan jatdingjiu tong **mzi** bin-go naamjan gitfan
Aafan must with MZI which-CL man marry
a. ✔ 'There is some man that Aafan must marry to.' scopally specific
b. ✘ 'Aafan must marry to a man (whoever he is).' free choice

A brief comparison with other languages:

- EIs in German, Italian Czech display non-uniform scope interaction with different intensional operators (Aloni and Port 2015; Šimík 2014);
- Obligatory wide scope indefinites are also attested in St'át'imcets (Matthewson 1999) and Tiwa (*-khi*, Dawson 2018), but *mzi*-indefinites can take *narrow* quantificational scope (see below).

2.3 Variable quantificational scope

Let us first focus on the indefinite meaning: it can scope above or below the universal quantifier.

(10) a. mui-go hoksaang dou hok-gwo [**mzi** bin-zung auzau jyujin]
every-CL student all learn-EXP MZI which-CL European language
'Every student has learned some European language.'

b. Wide: $\exists y[\text{an-unknown-European-language}(y) \wedge \forall x[\text{student}(x) \rightarrow \text{learned}(x,y)]]$

c. Narrow: $\forall x[\text{student}(x) \rightarrow \exists y[\text{an-unknown-European-language}(y) \wedge \text{learned}(x,y)]]$

Crucially, when interpreted narrowly, **the ignorance component is distributed** over 'every':
➜ For each student, s/he has learned some language unknown to the speaker.

Another example that favors a narrow scope reading:

(11) a. mui-sau 80nindoi coetman ge go dou hai goipin zi **mzi** bin-sau jatman-go
every-CL eighties famous GE song all be rearrange from MZI which-CL Japanese-song
'Every famous song in the eighties is rearranged from some Japanese song.'

b. #Wide: $\exists y[\text{an-unknown-Jap.-song}(y) \wedge \forall x[\text{a-famous-song}(x) \rightarrow \text{be.rearranged.from}(x,y)]]$

c. Narrow: $\forall x[\text{a-famous-song}(x) \rightarrow \exists y[\text{an-unknown-Jap.-song}(y) \wedge \text{be.rearranged.from}(x,y)]]$

➜ For each famous song in the 80's, it is rearranged from some Japanese song unknown to the speaker.

This contrasts with Japanese *-ka*, whose ignorance component disappears when interpreted narrowly.

(12) Japanese

Dono kyooju-mo dare-**ka** gakusee-to odotteru.
which professor-MO who-KA student-with is.dancing
'Every professor is dancing with some student.' (Alonso-Ovalle and Shimoyama 2014)

➜ it is felicitously continued by a follow-up question by the hearer: 'Who is dancing with who?'

3 Analysis

Taking stock, the ignorance component of *mzi*-indefinites shows a unique empirical profile:

(a) it cannot be cancelled or reinforced;
(b) it cannot be embedded under intensional operators (i.e. it is projective);
(c) it can take narrow quantificational scope and be distributed.

An overview of the proposal:

- (a) and (b) follows if the ignorance component is treated as **a conventional implicature**;
- (c) follows if *mzi* is **a choice function** that comes with this non-at-issue content.

3.1 Motivations for a conventional implicature approach

Adopting a general definition of conventional implicature, taken from **Potts:2015**, following Grice (1975) and Horn (2007),

(13) Meaning *p* is a *conventional implicature* of phrase S if, and only if:

a. *p* is a conventional (encoded) property of a lexical item or construction in S;

b. *p* is entailed by S; and,

c. *p*'s truth or falsity has no effect on the at-issue content of S.

The ignorance component of *mzi*-indefinites is obviously encoded by *mzi* and we have seen that it cannot be cancelled, satisfying (a) and (b).

Concerning (c) "*p*'s truth or falsity has no effect on the at-issue content of S", observe that the hearer can follow up by agreeing on the at-issue existential claim, while disputing the ignorance component:

(14) a. Aaming tai-zo [**mzi** bin-bun syu] =(3)
Aaming read-PERF MZI which-CL book
'Aaming read some book (I don't know which).'

b. hai aa3. batgwo nei jinggoi zidou hai bin-bun gaa3
yes SFP but you probably know be which-CL SFP
'Yes, but you probably know which book it is.'

(cf. Karttunen and Peters 1979; Potts 2005)

Note that (14b) is an infelicitous follow-up to (15).

(15) ngo **m-zi** [Aaming tai-zo bin-bun syu] =(4)
I not-know Aaming read-PERF which-CL book
'I don't know which book Aaming read.'

3.2 Motivations for a choice-functional analysis

Mzi-indefinties display 'exceptional wide scope', taking scope from within a syntactic island.

(16) mui-go hoksaang dou tengdou [hokhaau kwaidingjiu hok [**mzi** bin-zung jyujin] ge siusik]
every-CL student all heard school require learn MZI which-CL language GE news
'There is some language s.t. every student heard the news that the school requires (them) to learn it.'
(A narrow scope reading is less salient but possible.)

I therefore adopt a choice-functional approach to *mzi*-indefinites, following Kratzer (1998), Reinhart (1997), and Winter (1997).

3.3 Implementation

(17) A multi-dimensional semantics of *mzi* (Karttunen and Peters 1979; Potts 2005)

a. *At-issue content*:
$[\![mzi_i]\!]^g = \lambda P_{<e,t>}.\ g(i)(P)$, where $g(i) \in D_{\text{choice function } <<e,t>,e>}$

b. *Conventional implicature*:
The speaker *doesn't know* (i.e. fails to identity in a relevant way) the referent chosen by the choice function.

An illustration: to derive the narrow scope reading of (18a), with the distributed ignorance component

(18) a. mui-go hoksaang dou hok-gwo [**mzi** bin-zung auzau jyujin] =(10)
every-CL student all learn-EXP MZI which-CL European language
'Every student has learned some European language.'

b. Narrow: ∀x[student(x) → ∃y[an-unknown-European-language(y) ∧ learned(x,y)]]

Assuming that *wh*-expressions denote alternative sets (Kratzer and Shimoyama 2002; Beck 2006, i.a.),

(19) a. The internal structure of the *mzi*-indefinites:
[$_{NP}$ *mzi* [$_{NP}$ which European.language]]

b. *At-issue-content*: $[\![mzi_i]\!]^g$ ([[which European.language]]) via (17a))
= λX. g(i)(X) ({x: European.language(x) }) by Functional Application
= g(i){x: European.language(x) }
= g(i){Spanish, German, ... }

Here, I assume that the choice function is existentially bound at its base position (Winter 1997).

(20) The meaning of (18a)

a. *At-issue-content*: ∀x[student(x) → ∃f [learned(x, f{Spanish, German, … })]]

b. *Conventional implicature*: The speaker doesn't know the referent chosen by f.

Since the ignorance component is associated with the choice function, it is distributed altogether.

4 Evidence for grammaticalization

Returning to the source of the ignorance component in *m-zi*, I suggest:

- the (negated) attitude verb *m-zi* grammaticalizes syntactically as **an adnominal modifier** and semantically as **a choice function**.
- Instead of contributing to the at-issue content, its lexical meaning is carried over to the choice function, which then denotes a specific type of choice function.

➜ Since *mzi* has a different origin compared to EIs in other languages, it is not surpring that it has a different empirical profile.

4.1 Three positions of *m + zi*

The grammaticalization path: ❶ an attitude verb ➜ ❷ a raising verb ➜ ❸ a choice function.

(21) ngo **m-zi** [Aaming tai-zo bin-bun syu] an attitude verb; =(4)
I not-know Aaming read-PERF which-CL book
'I don't know which book Aaming read.'

(22) Aaming **m-zi** tai-zo bin-bun syu a "raising" verb
Aaming not-know read-PERF which-CL book
'It is not known which book Aaming read.' (cf. "an attitudinal marker", Yap and Chor 2014)

(23) Aaming tai-zo [**mzi** bin-bun syu] an adnominal modifier; =(3)
Aaming read-PERF MZI which-CL book
'Aaming read some book (I don't know which).'

m-zi as ...	Attitude holder	Complement	Ignorance
❶ an attitude verb	overt	clauses	at-issue
❷ a raising verb	covert	interrogative clauses	at-issue
❸ an adnominal modifier	covert	WH	non-at-issue

Table 1: Different usages of the string *m-zi*

(Note that *m-zi* can also take nominal complements, not included in the table.)

4.2 Corpus data

With reference to two corpora:

(i) Early Cantonese Colloquial Texts: A Database (data mainly in 19th century) and;
(ii) A Linguistic Corpus of Mid-20th Century Hong Kong Cantonese

m-zi as ...	(i) Early Can.	(ii) Mid-20th HKC
❶ an attitude verb	4/60	44/110*
❷ a raising verb	2/60	17/110*
❸ an adnominal modifier	0/60	7/110*

Table 2: Frequency of *m-zi* (*total hit: 1098, counting the first 10%)

(The counting omits instances of *m-zi* in A-not-A form, in answer fragments, in idioms, or with null/nominal arguments.)

- Compared to ❷, ❸ emerges relatively recently.
- The usage of ❷ is more frequent than ❸.

4.3 Fusion of predicate and *wh*-expressions

Cross-linguistic data reveal that it is not uncommon for *wh*-expressions to develop into indefinites by fusing with predicates (Haspelmath 1997, p.131).

(24) a. Middle High German
ne weil wer '(I) don't know who' ➜ *neizwer* 'somebody'

b. Old English
ne wät hwā '(I) don't know who' ➜ *näthwä* 'somebody'

c. French
Je ne sais (pas) quel 'I don't know which' ➜ *je ne sais quel* 'some kind of'

5 Conclusions

In this talk, I have showed that:

- Cantonese EIs have a different morphological makeup than other more discussed EIs.
- The nature : is **a conventional complicature** (Grice 1975; Potts 2005; Horn 2007), representing a third type of the ignorance component (different from ❶ and ❷);
- The source : it originates from the lexical meaning of *m-zi* 'not-know' and becomes a non-at-issue content due to grammaticalization of *m-zi* into **a choice function**.

A comparison on how languages bundle the existential claim and the ignorance inference:

	Cantonese	Spanish	Japanese	Tiwa
Form	*mzi* + WH	*algún* + NP	WH + *ka*	WH + *khi*
Cancellability	No	Yes	Yes	No
Intensional scope	Wide	Wide/narrow	Wide/narrow	Wide
Quantificational scope	Wide/narrow	Wide/narrow	Wide/narrow	Wide
Distributed ignorance	Yes	No	No	N/A
Nature	Conventional imp.	Quantity imp.	Quantity imp.	Manner imp.

Some further issues:

- to what extent the properties of the ignorance component reveal how EIs emerge (e.g. grammaticalization, lexical competition, conventionalized conversational implicature, etc.)
- why a language adopts a particular way of bundling, but not the other

Appendix A: Methods of identification

Mzi is compatible with ignorance concerning different identification methods.

(25) a. Aaming zinghai dak **mzi** bin-bun syu soeng maai
Aaming only only MZI which-CL book want buy
'Aaming wants to buy only some book.' ✔ description; ✘ naming/ostension

b. Aaming tai-zo **mzi** bin-bun giuzou Hunglaumung ge syu
Aaming read-PERF MZI which-CL titled Dream.of.the.Red.Chamber GE book
'Aaming read some book titled *Dream of the Red Chamber*.'
✔ naming; ✘ ostension/descprtion

It differs from EIs in Romance languages (where there is an *otension>naming>description* hierarchy, Aloni and Port 2015), and Sinhala (where the epistemic marker specifies the unknown methods of identification, Slade 2015)

References

Aloni, Maria, and Angelika Port. 2015. "Epistemic indefinites and methods of identification." In *Epistemic Indefinites,* edited by Luis Alonso-Ovalle and Paula Menéndez-Benito, 117–140. Oxford: Oxford University Press.

Alonso-Ovalle, Luis, and Paula Menéndez-Benito. 2010. "Modal indefinites." *Natural Language Semantics* 18 (1): 1–31.

———. 2013. "Two Views on Epistemic Indefinites." *Linguistics and Language Compass* 7 (2): 105–122.

———. 2015. *Epistemic Indefinites.* Oxford: Oxford University Press.

Alonso-Ovalle, Luis, and Junko Shimoyama. 2014. "Expressing Ignorance in the Nominal Domain: Japanese *Wh*-ka." In *Proceedings of the 31st West Coast Conference on Formal Linguistics,* edited by Robert E. Santana-LaBarge, 11–20. Somerville, MA: Cascadilla Proceedings Project.

Beck, Sigrid. 2006. "Intervention effects follow from focus interpretation." *Natural Language Semantics* 14:1–56.

Dawson, Virginia. 2018. "A new kind of epistemic indefinite." In *Proceedings of Sinn und Bedeutung,* edited by Uli Sauerland and Stephanie Solt, 22:349–366. Berlin: ZAS.

Geist, Ljudmila. 2008. "Specificity as Referential Anchoring : Evidence from Russian." In *Sinn und Bedeutung,* edited by Atle Grønn, 151–164. Oslo: ILOS.

Giannakidou, Anastasia, and Josep Quer. 2013. "Exhaustive and non-exhaustive variation with free choice and referential vagueness: Evidence from Greek , Catalan , and Spanish." *Lingua* 126:120–149.

Grice, H. P. 1975. "Logic and conversation." In *Syntax and semantics 3: Speech arts,* edited by Peter Cole and Jerry L Morgan, 41–58. New York: Academic Press.

Haspelmath, Martin. 1997. *Indefinite Pronouns.* Oxford: Oxford University Press.

Horn, Laurence R. 1972. "On the semmantic properties of logical operators in English." PhD diss., University of California, Los Angeles.

———. 2007. "Toward a Fregean pragmatics : Voraussetzung, Nebengedanke, Andeutung." In *Explorations in Pragmatics : Linguistics, cognitive and intercultural aspects,* edited by Istvan Kecskes and Laurence R. Horn, 39–69. Berlin: Mouton de Gruyter.

Jayez, Jacques, and Lucia M. Tovena. 2006. "Epistemic Determiners." *Journal of Semantics* 23:217–250.

Karttunen, Lauri, and Stanley Peters. 1979. "Conventional implicature." In *Syntax and semantics, xi : Presupposition,* edited by Choon-kyu Oh and David A Dinneen, 1–56. New York: Academic Press.

Kratzer, Angelika. 1998. "Scope or pseudoscope? Are there wide-scope indefinites?" In *Events and grammar,* edited by Susan Rothstein, 163–196. Dordrecht: Kluwer Academic Publishers.

Kratzer, Angelika, and Junko Shimoyama. 2002. "Indeterminate pronouns: The view from Japanese." In *Proceedings of the Tokyo conference on psycholinguistics,* edited by Yukio Otsu, 3:1–25. Tokyo: Hituzi Syobo.

Matthewson, Lisa. 1999. "On the Interpretation of Wide-Scope Indefinites." *Natural Language Semantics* 7 (1): 79–134.

McCawley, James D. 1978. "What is conversationally implicated by an utterance depends not only on the utterance but on what other utterances the speaker could have produced but did not. For example , a declarative sentence A." In *Syntax and Semantics 9: Pragmatics,* edited by Peter Cole, 245–259. Academmic Press.

Potts, Christopher. 2005. *The Logic of Conventional Implicatures.* Oxford: Oxford University Press.

Reinhart, Tanya. 1997. "Quantifier Scope: How Labor is Divided between QR and Choice Functions." *Linguistics and Philosophy* 20 (4): 335–397.

Sadock, Jerrold M. 1978. "On testing for conversational implicature.pdf." In *Syntax and Semantics 9: Pragmatics,* edited by Peter Cole, 281–297. Academic Press.

Šimík, Radek. 2014. "Epistemic indefinites under epistemic modals in Czech." In *Slavic grammar from a formal perspective: Proceedings of FDSL 10,* edited by Gerhild Zybatow, Petr Biskup, Marcel Guhl, Claudia Hurtig, Olav Mueller-reichau, and Maria Yastrebova, 425–442. Frankfurt am Main: Peter Lang.

Slade, Benjamin. 2015. "Sinhala epistemic indefinites with a certain *je ne sais quoi*." In *Epistemic Indefinites: exploring modality beyond the verbal domain,* 82–99. Oxford: Oxford University Press.

Winter, Yoad. 1997. "Choice Functions and the Scopal Semantics of Indefinites." *Linguistics and Philosophy* 20 (4): 399–467.

Yap, Foong Ha, and Winnie Oi-Wan Chor. 2014. "Epistemic, evidential and attitudinal markers in clause-medial position in Cantonese." In *Modes of Modality: Modality, typology, and universal grammar,* edited by Elisabeth Leiss and Werner Abraham, 219–262. Amsterdam and Philadelphia: John Benjamins Publishing Company.