Epistemic indefinites in Cantonese:

a case study of *m-zi* 'not-know'*

Tommy Tsz-Ming Lee

City University of Hong Kong

August 7, 2023

Abstract

A recent line of research (Alonso-Ovalle and Menéndez-Benito 2015, and references therein) concerns how an existential claim is bundled with a modal component — in particular, one relating to speakers' ignorance — in the nominal domain . The primary focus of this paper is to discuss a case of epistemic indefinites in Cantonese which takes the form of *m-zi-WH* 'not-know-WH'. I argue for two claims in this paper. First, I propose that *mzi* is an (overt) choice function operator that binds *wh*-expressions and disjunctive expressions. Second, I suggest that the ignorance component associated with *mzi* is a conventional implicature, representing a new type of the ignorance component for epistemic indefinites. I also discuss a potential grammaticalization path of *m-zi*, connecting its syntactic distribution to its semantic properties. The findings in this paper uncover a new breed of epistemic indefinites and contributes to our understanding of both the interpretation of *wh*-expressions in Cantonese and the study of how natural languages encode an

ignorance component in nominal expressions.

1 Introduction

A recent line of research concerns how a particular existential claim and a modal component related to speaker's ignorance are bundled in the nominal domain. The relevant expressions are often referred to as *modal indefinites* or *epistemic indefinites* (see Alonso-Ovalle and Menéndez-Benito 2015, for recent overview and references therein). The primary focus of this paper is to discuss a case of epistemic indefinites in Cantonese, whose major differences from epistemic indefinites in other languages lie in the less common morphological makeup, and the properties of the associated ignorance component.

The Cantonese expression of interest is the string *m-zi*, which consists of a negative (bound) morpheme *m* 'not' and the verb *zi* 'know'. Canonically, *m-zi* is construed as an attitude verb, taking either a declarative complement (as in (1a)) or an interrogative complement (as in (1b)).[1]

(1) a. ngo **m-zi** [Aaming tai-zo   ni-bun syu]   declarative complement

       I   not-know Aaming read-PERF this-CL book

       'I don't know Aaming read this book.'

   b. ngo **m-zi** [Aaming tai-zo   bin-bun syu]   interrogative complement

       I   not-know Aaming read-PERF which-CL book

       'I don't know which book Aaming read.'

The focus of this paper is a less discussed usage of *m-zi*, where it forms a constituent with a *wh*-expression. For example, in (2), the expressions containing *mzi* and a *wh*-expression serves as the object of the sentence.[2] [3]

---

1. *M-zi* can also take nominal complements, as in (i). I set aside this usage throughout this paper.

(i) ngo **m-zi**   ni-gin si
   I   not-know this-CL event
   'I don't know this.'

2. Anticipating the discussions in §2.1, I transcribe this usage of *m-zi* as *mzi* and gloss it with MZI to distinguish it from its usage as an attitude verb.

3. At first glance, the examples in (2) seem to correspond to English *John read I don't know which book* and *John gave a book to I don't know which student*. But as will be clear in §2.1.2, the usage of *mzi* is substantially different from English *don't know* in these sentences.

(2) a. Aaming tai-zo   [**mzi** bin-bun   syu]   as direct object

Aaming read-PERF MZI   which-CL book

⇝ 'Aaming read some book (I don't know which).'

b. Aaming bei-zo   jat-bun syu [**mzi** bin-go   hoksaang]   as indirect object

Aaming give-PERF one-CL book MZI   which-CL student

⇝ 'Aaming gave a book to some student (I don't know which student).'

This usage of *mzi* is interesting in two ways. First, the *mzi+wh* expressions come with an ignorance component associated with the speaker, i.e., the speaker conveys his/her inability to identify the referent. This property is characteristic of *epistemic indefinites*, which have attracted much attention in the literature (Kratzer and Shimoyama 2002; Jayez and Tovena 2006; Sudo 2010; Alonso-Ovalle and Menéndez-Benito 2010; Aloni and Port 2015; Slade 2015; Chen et al. 2017; Dawson 2018; Liu and Yang 2021, i.a.). Two examples are given with the indefinite markers in Spanish, i.e., *algún*, and in German, i.e., *irgendein*, below. The signature property of epistemic indefinites is their incompatibility with continuations that presume the speaker's knowledge.

(3) a. María se casó   con #**algún**/ un estudiante del   departamento de lingüística:

María SE married with ALGÚN/ UN student   of.the department   of linguistics

en concreto con   Pedro.

namely   with Pedro

'María married a linguistics student, namely Pedro.'

(Spanish, Alonso-Ovalle and Menéndez-Benito 2010, p.2)

b. **Irgendein**   Student hat angerufen. #Rat   mal wer?

IRGEND-one student has called   guess PRT who

'Some student called. Guess who?'   (German, Aloni and Port 2015, p.117)

However, the nature of the ignorance component in different languages appears to be non-uniform (see, for an overview on different approaches, Alonso-Ovalle and Menéndez-Benito 2013). In some Gricean approaches, it is a quantity implicature (Alonso-Ovalle and Menéndez-Benito 2010; Alonso-

Ovalle and Shimoyama 2014; Chierchia 2013; Chierchia and Liao 2015; Dawson 2018). In other approaches, it is taken to be a consequence of a shift on the method of identification required by the context (Aloni and Port 2015; Šimík 2014; Slade 2015). The question for *mzi+wh* in Cantonese is then: (i) is it an instance of epistemic indefinites, given its less common morphological makeup[4], and (ii) if it is, what is the nature of the ignorance component?

Another issue of interest concerns the interpretation of *wh*-expressions in Cantonese. Both sentences in (2) are interpreted as declarative sentences but not matrix questions. The *wh*-expressions are thus not used interrogatively but convey an existential meaning when combined with *mzi*. While the existential reading of *wh*-expressions has been widely discussed in Mandarin (Chinese), it is not so in Cantonese. It has long been noticed that bare *wh*-expressions in Mandarin can be interpreted existentially in modalized contexts, as in (4) (see Li 1992; Lin 1998; Chierchia and Liao 2015, i.a.; see also Chen et al. 2017; Liu and Cui 2019; Liu and Yang 2021 for more recent discussions on the ignorance component).

(4) Zhangsan *(haoxiang) chi-le   shenme   Mandarin

Zhangsan seemingly   eat-PERF   what

'Zhangsan seemingly ate something.'

However, bare *wh*-expressions in Cantonese lack such an existential reading even in modalized contexts, as illustrated in (5), revealing a difference between Cantonese and Mandarin in terms of the licensing of the existential interpretations of *wh*-expressions.[5]

---

4. Many cases of epistemic indefinites in Romance and Germanic languages involve articles/determiners (see, e.g., Kratzer and Shimoyama 2002; Jayez and Tovena 2006; Alonso-Ovalle and Menéndez-Benito 2010); in article-less languages, they are often marked by question particles such as Japanese (Sudo 2010; Alonso-Ovalle and Shimoyama 2014) and Sinhala (Kishimoto 2005; Slade 2015). But as we will see in §5, Cantonese is not unique in employing a verb as the epistemic marker.

5. While I do not aim to provide an explanation as to why bare *wh*-expressions in Cantonese cannot be used existentially, the observation is in line with the suggestion in Haspelmath (1997, p.176) that "the interrogative function [of *wh*-expressions] is always primary, and the indefinite function is secondary" and that "[a]n interrogative pronoun may lose its indefinite function" (e.g. Latin *quis/quem* 'who(m)' and Classical Greek *tís* 'who'). This describes the Cantonese case.

(5) Aaming (houci)   sik-zo   matje   Cantonese

    Aaming seemingly eat-PERF what

    Interrogative: 'What did Aaming (seemingly) eat? '

    $\neq$ Existential: 'Aaming (seemingly) ate something.'

The expression *mzi+wh* thus represents an understudied case of the existential interpretation of *wh*-expressions in Cantonese.[6] The question is: (i) what is the semantic contribution of *mzi*, and (ii) how does it interact with *wh*-expressions?

Against this background, I argue for two proposals in this paper. First, I propose that *mzi* is an (overt) choice function operator (Kratzer 1998; Winter 1997; Reinhart 1997) – it takes an alternative set as its argument and returns a member of the set. It can bind *wh*-expressions and disjunctive expressions and contributes an existential meaning. Second, I suggest that the ignorance component associated with *mzi* is a conventional implicature (in the sense of Karttunen and Peters 1979; Potts 2005), representing a new type of the ignorance component for epistemic indefinites. Building on these two proposals, I further discuss a potential grammaticalization path of *m-zi*, connecting its syntactic distribution to its semantic properties. This paper thereby uncovers a new breed of epistemic indefinites, and contributes to our understanding of both the interpretation of *wh*-expressions in Cantonese and the study of how natural languages encode an ignorance component in nominal expressions.

The rest of this paper is organized as follows. §2 differentiates the non-verbal usage of *mzi* from its verbal usage. I present arguments in favor of a choice-functional analysis of *mzi*. §3 specifically discusses the ignorance component associated with *mzi*. I examine its interaction with the quantificational domain, intensional operators, quantifiers and negation. I suggest that the ignorance component is best regarded as a conventional implicature. Building on the findings in §2 and §3, I sketch a compositional analysis on *mzi* in §4. In §5, I discuss a potential grammaticalization path of *mzi*. I

---

6. There are two other cases where *wh*-expressions are reported to be interpreted existentially. One is in its doublet form, e.g. *matje-matje* 'what-what', see Wong (2018) and Lee and Wong (2018). Another is *wh*-placeholder, which take the form of demonstrative-classifier-*wh*, e.g., *go-go matje* "(lit.) that what"; see Cheung (2015) for discussions on its metalinguistic usage in Mandarin. As far as I can tell, Cantonese displays a similar usage.

conclude in §6.

2   A choice functional analysis of *mzi*

In §2.1, I suggest that *mzi* as used in (2) differs substantially from with its verbal use as in (1). I then argue that *mzi* cannot be treated as a type of syntactic amalgam that blends two sentences into one. In §2.2, I provide three pieces of evidence for a choice-functional analysis of *mzi*. I show that *mzi* selects as its associate expressions that invoke sets of alternatives for their evaluation. These expressions include *wh*-expressions and disjunctive phrases. Additionally, *mzi* competes with other potential binders of *wh*-expressions, and thus should be treated on a par with other *wh*-binders. Lastly, I illustrate that when *mzi* forms an indefinite with a *wh*-nominal, it displays exceptional wide scope behaviors.

2.1   The non-verbal use of *mzi*

2.1.1   Differentiating the verbal *m-zi* from the non-verbal *mzi*

To differentiate the two usages of *m-zi/mzi*, it is crucial to show that the target cases of *mzi* involve a nominal structure instead of a clausal one. This is already shown in (2a), where *mzi* forms a constituent with a *wh*-expression, repeated in (6). This constituent serves as the argument of the transitive verb. Since the verb *tai* 'read' only selects a nominal complement (not a verbal/clausal one), the *mzi+wh* expression in the post-verbal position indicates its nominal nature.

(6)   Aaming tai-zo   [**mzi** bin-bun   syu]   =(2a)
      Aaming read-PERF MZI   which-CL book
      ⇝ 'Aaming read some book (I don't know which).'

Coordination test illustrates the same point. The *mzi-wh* string can conjoin with a plain indefinite, as in (7). It conveys a meaning where Aaming read multiple things - one is a magazine, another is a book (but the speaker does not know which book it is). Such reading is unexpected if *mzi* is a verb and

forms a VP conjunction with the verb phrase 'read a magazine' (which means something like 'Aaming read a magazine and didn't know some book').[7]

(7) Aaming tai-zo   [jat-bun zaapzi   tungmaai **mzi** bin-bun   syu]

Aaming read-PERF one-CL   magazine and   MZI which-CL book

'Aaming read [a magazine and some book].'

As far as meaning is concerned, there is a subtle but detectable difference between (1b) and (2a)/ (6). (1b) is repeated below.[8]

(8) ngo **m-zi**   [Aaming tai-zo   bin-bun   syu]   =(1b)

I   not-know Aaming read-PERF which-CL book

'I don't know which book Aaming read.'

At first glance, the two sentences convey very similar meaning, but they can be contrasted in terms of whether the existence of the book that Aaming read is *presupposed* or *asserted*. For example, while (8) is compatible with the continuation in (9), (6) is not.

(9) A felicitous continuation for (8) but not (6)

daan honeng keoi bin-bun   dou mou   tai-gwo

but   perhaps he   which-CL all   not.have read-EXP

'But perhaps he didn't read any book.'

Since the continuation in (9) denies the existence of the book that Aaming read, the contrast reflects that in (1b) the existence is not asserted, but just presupposed (as in interrogative clauses). Thus it can be retreated by uttering (9). However, (2a) *asserts* such existence. Denying such existence in the continuation immediately contradicts (2a), hence there is infelicity. I take this as an important semantic distinction between the predicative *m-zi* and the non-predicative *mzi*.[9]

---

7. For the rest of the paper, I will translate *mzi* as "some", without explicitly mentioning the ignorance component in the translation for convenience.

8. I thank Sze-Wing Tang for raising questions on the meaning difference between these two sentences.

9. Anticipating a choice functional analysis on *mzi*, the existential quantification comes from an existential quantifier associated with *mzi*.

2.1.2   Not syntactic amalgams

Before we conclude the non-verbal status of *mzi*, one alternative has to be taken into consideration. The usage of *mzi* in (2) is reminiscent of a type of syntactic amalgam in English, particularly, the Andrews Case, as discussed in Lakoff (1974). Some examples are given below:

(10) a.   Sally will eat <u>I don't know what</u> today.   (Johnson 2013, p.73)

     b.   Sammy's going to marry <u>guess who</u>.   (Lakoff 1974, p.321)

     c.   John is going to marry <u>what's her name</u>?   (p.c. Deniz Rudin)

The signature property of Andrews amalgams is that the two clauses are amalgamated together and the *wh*-expression simultaneously serves as an argument of both the matrix verb and the verb in the (underlined) interrupting clause. One way to maintain the verbal status of *mzi* is to suggest that the Cantonese sentences in (2) are also resulted from syntactic amalgams and the *wh*-expression is selected by both the matrix verb and *mzi*.

     However, what we observe so far for *mzi* is that the position it occupies is highly restricted to just one verb. As shown in (11), no other verb can replace *mzi*.

(11)   Aaming tai-zo   { ??**m-geidak**/   ***m-hangding**/ ***mou-lamhei**}   bin-bun   syu
      Aaming read-PERF   not-remember/ not-be.sure/   not.have-think.of which-CL book
      'Aaming read some book that { I do not remember/ I am not-sure/ I didn't think of }.'

The Cantonese case of *mzi* thus contrasts with the examples in (10), where syntactic amalgams are observed with a wide range of verbs (e.g., *know*, *guess*, etc.) and constructions (e.g., declaratives, imperatives, questions, etc.). This suggests some idiosyncratic development of the string *mzi*, instead of a more general syntactic mechanism that includes *mzi* (see also discussions in §5).

     Also, the string *mzi* must appear as such without any alternation, i.e., it behaves like a single unit without internal structure. For example, it is not possible to omit the negation, or insert a verb between *m* and *zi*, as in (12). This further indicates that *mzi* might have undergo lexicalization and become a single lexical item.

(12) *Aaming tai-zo   { **zi**/   **m-soeng-zi**}   bin-bun   syu

     Aaming read-PERF   know/ not-want-know which-CL book

     (Lit.) 'Aaming read { know/ not-want-know } which book.'

Furthermore, if *mzi* were a verb, we would expect recovering the subject (i.e., the attitude holder) in the *mzi-wh* string to be possible. The inability to do so, however, suggests that the subject position of *mzi* is unavailable and thus the *mzi*-WH string does not form a clause.

(13) *Aaming tai-zo   ngo **mzi** bin-bun   syu

     Aaming read-PERF I   MZI   which-CL book

     'Aaming read I don't know which book.'

Lastly, the string *mzi bin-bun syu* 'MZI which book' by itself is degraded, as in (14a) and a copula before the *wh*-expression is preferred, as in (14b). However, with the presence of the copula, (14b) cannot replace the string *mzi bin-bun syu* in (2a). These observations combine to suggest that *mzi* is substantially different from its verbal counterpart *m-zi* 'not-know'.

(14) a. ??(ngo) **m-zi**   bin-bun   syu   b. (ngo) **m-zi**   hai bin-bun   syu

         I   not-know which-CL book   I   not-know be which-CL book

         '(I) don't know which book.'   '(I) don't know which book it is.'

In short, this subsection showed that *mzi* can exceptionally occupy positions that are unavailable for other verbs. I conclude the non-verbal status of *mzi* as used in (2). In the next subsection, I present arguments for a choice-functional analysis of *mzi*.

2.2 Arguments for a choice-functional analysis of *mzi*

2.2.1 *Mzi* and expressions invoking sets of alternatives for evaluation

The first argument concerns a semantic restriction on the associate of *mzi*. *Mzi* forms a constituent not only with *wh*-nominals, but also with *wh*-adverbials (conveying a non-interrogative reading). The

mzi+*wh*-adverbial string occupies the same position as other (non-*wh*) adverbials. Observe the following examples:

(15) a. Aaming houci { **mzi** dimjoeng/ houfaaigam } zinghou-zo gaa-ce

Aaming seem   MZI how   rapidly   fix-PERF   CL-car

'Aaming seems to have { somehow/ rapidly} fixed the car.'

 b. Aaming camjat   { **mzi** dimgaai/ janwai ze   syu } lai-zo

Aaming yesterday   MZI why   because borrow book   come-PERF

'Aaming came yesterday { for some reason/ for book-borrowing }.'

In addition to *wh*-expressions, *mzi* can combine with disjunctive phrases, and these disjunctive phrases can be of different syntactic size, including nominal, verb phrases and clauses. Note that disjunctive marker must be the one for alternative questions (i.e., *ding* "or"), but not the one for logical, boolean disjunction (i.e., *waakze* "or").[10]

(16) a. NP disjunction

Aaming maai-zo **mzi** [NP syu   {ding/ *waakze} bouzi]

Aaming buy-PERF MZI   book or/   or   newspaper

'Aaming bought some book or newspaper.'

 b. VP disjunction

Aaming zoeng ni-bun syu   **mzi** [VP fong-zo   hai toi   {ding/ *waakze}

Aaming take   this-CL book MZI   put-PERF on table or/   or

 daai-zo-faan   hokhaau]

bring-PERF-back school

'Aaming either put this book on the table or brought it back to school.'

 c. Sentential disjunction

---

10. Whether the formation of disjunctive phrases involves joining different phrases or are derived via Conjunction Reduction does not concern us here. See, for discussions of their Mandarin counterpart, **HuangR:2010**; Erlewine (2017).

Aaming munggin **mzi** [CP/TP keoi saat jan {ding/ *waakze} jau jan saat

Aaming dream MZI he kill person or/ or have person kill

keoi]

he

'Aaming dreamed that either he killed someone or someone killed him.'

Relatedly, these observations suggest that *mzi* preserves the syntactic categories of its associate - the presence of *mzi* does not alter the category of its associate. In all the sentences in (15) and (16), leaving out *mzi* still results in well-formed sentences (but they become matrix questions).

The range of possible grammatical elements that can be associated with *mzi* is not entirely unconstrained. *Mzi* cannot combine with referential expressions such as definite/indefinite expressions and proper names.

(17) *Aaming tai-zo **mzi** { ni-bun syu/ jat-bun syu/ Hunglaumung}

Aaming read-PERF MZI this-CL book/ one-CL book/ Dream.of.the.Red.Chamber

(Lit.) 'Aaming read { this book/ one book/ Dream of the Red Chamber }that I don't know.'

In other words, expressions combined with *mzi* must be either *wh*-expressions or interrogative disjunctive phrases (marked by *ding*). These elements can be regarded as a natural semantic class if we assume that both *wh*-expressions and disjunctive phrases invoke sets of alternatives for their evaluation (see, for the former, Hamblin (1973); for the latter, Simons (2005), Alonso-Ovalle (2006), and Aloni (2007)). However, the observation that indefinites are incompatible with *mzi* (which arguably denote alternatives in the neo-Hamblin semantics (Kratzer and Shimoyama 2002; Charlow 2019)) suggests that invoking alternatives is only a necessary condition. I suggest that the associate of *mzi* must also be interrogative by nature (which excludes non-interrogative disjunctive phrases marked by *waakze* 'or'). We thus reach the following requirement of *mzi* on its associate.

(18) The syntax-semantic requirement of *mzi* on its associate

The associate of *mzi* must be an interrogative expression that invokes sets of alternatives for their evaluation.

Together with the existential meaning contributed by *mzi* (as in (2), I suggest that it is best regarded as a choice function in the sense of Kratzer (1998), Winter (1997), and Reinhart (1997), which takes a set of alternatives as its argument and returns a member of that set (for a compositional analysis, see §4).[11]

2.2.2 Competitions with other *wh*-binders

The second argument concerns competitions with other *wh*-binders. As with Mandarin and many other languages, the interpretation of *wh*-expressions in Cantonese is dependent on their binders (Kratzer and Shimoyama 2002; Shimoyama 2006; Chierchia and Liao 2015, i.a.). The following examples show that *wh*-expressions in Cantonese can (i) have an interrogative meaning when associated with some covert question operator in (19a), (ii) serve as negative polarity items under negation (19b), or (iii) give rise to a universal reading when co-occurring with the universal/distributive quantifier *dou* in (19c).

(19)   a.   Interrogative (⤳ covert question operator)

        Aaming sik-zo   matje?

        Aaming eat-PERF what

        'What did Aaming eat?'

     b.   Negative polarity item *any* (⤳ negation)

        Aaming mou sik matje   (Lee 2014)

        Aaming not   eat what

        'Aaming didn't eat anything.'

     c.   Universal (⤳ *DOU*-quantifier)

        Aaming matje dou sik

        Aaming what all   eat

        'Aaming eats everything.'

---

11. The incompatibility of *mzi* with bare indefinites does not follow from a choice-functional analysis; see §5 for how *mzi* inherits the interrogative requirement from its predicative usages.

In this regard, *mzi* pattern with these binders of *wh*-expressions in that *mzi* can also combine with various *wh*-expressions, as shown in Table 1 below.

Base gloss	*mzi*	Interrogative	Negative *mou*	Universal *dou*
what	*mzi matje*	*matje*	*mou...matje*	*matje dou*
who	*mzi bingo*	*bingo*	*mou...bingo*	*bingo dou*
which	*mzi bin-*CL	*bin-*CL	*mou...bin-*CL	*bin-*CL *dou*
how$_{manner}$	*mzi dimjoeng*	*dimjoeng*	*mou...dimjoeng*	*dimjoeng dou*
how many	*mzi gei do*	*gei do*	*mou...gei do*	*gei do dou*
when	*mzi geisi*	*geisi*	-	*geisi dou*
why	*mzi dimgai*	*dimgai*	-	-

Table 1: Possible binders of *wh*-expressions in Cantonese

If *mzi* is also a *wh*-binder, we expect to see that different binders may compete for the binding relation with a *wh*-expression. We have already seen an example where the presence of *mzi* blocks an interrogative reading - the *wh*-expressions in (2) can only be interpreted existentially. The absence of an interrogative reading can be explained by the fact that the *wh*-expressions are bound by the lower *mzi* before the higher, clause-level interrogative operator.

The same can be said with regard to embedded questions. In an embedded clause containing a *wh*-expression, if *mzi* is present and forms a constituent with the *wh*-expression, then the embedded clause fails to serve as an interrogative complement to *soeng-zidou* 'wonder (lit. want-know)', as in (20).

(20) *ngo soeng-zidou [Aaming tai-zo **mzi** bin-bun syu]

      I   wonder   Aaming read-PERF MZI which-CL book

      Int.: 'I wonder what book Aaming read.'

Similar blocking effects are observed with negation and the universal quantifier *dou*. In (21a), the presence of *mzi* in (21a) forces an existential reading (as it is closer to the *wh*-expression). Also, in (21b), the universal reading is unavailable in the presence of *mzi*. The *wh*-expression in (21b) is construed existentially.

(21) a. Aaming mou   tai **mzi** matje

Aaming not.have read ᴍᴢɪ what

'There is something that Aaming didn't read.'

≠ 'Aaming didn't read anything.'

b. **mzi** bin-bun syu   dou maai-saai

ᴍᴢɪ which-ᴄʟ book all   sold.out

'There is a book sold out (but I don't know which).'

≠ 'Every book is sold out.'

In sum, *mzi* shares distributional similarities with other *wh*-binders and it blocks the binding relation between a *wh*-expression and another potential binder. I take this as evidence for the binder/operator status of *mzi*.

2.2.3   Indefiniteness and scope behaviors of the *mzi+wh* string

The last argument concerns the scope behaviors of the *mzi+wh* string. Taking *mzi* as a choice function operator that binds a *wh*-expression leads us to expect that *mzi+wh* strings display canonical properties of indefinites. This prediction is borne out in a number of ways.

First, the *mzi+wh* string is compatible with existential *jau*-constructions. The construction arguably displays a *Definiteness Effect* (Safir 1982; Huang 1987), requiring that the nominal following *jau* 'have' must be indefinite. The *mzi+wh* string occurs with plain indefinites but not definite expressions.

(22) jau   {**mzi** bin-bun   syu/   jat-bun syu/   *ni-bun syu} mgin-zo

have ᴍᴢɪ   which-ᴄʟ book/ one-ᴄʟ book/ this-ᴄʟ   book lost-ᴘᴇʀꜰ

'There is {some book/one book/ this book} that is lost.'

Another piece of evidence comes from the minimal pairs below. The sentence in (23a) constitutes a contradiction when a definite expression is used (i.e., Aaming). In contrast to this, the *mzi-wh* string does not give rise to any contradictory reading in (23b), suggesting its indefinite status (see also a

similar test on Tiwa in Dawson (2018)).

(23) a. #Aaming lai-zo,   Aaming mou lai

    Aaming come-PERF Aaming not   come

    'Aaming came, Aaming didn't come.'

   b. **mzi** bingo lai-zo,   **mzi** bingo mou lai

    MZI who   come-PERF MZI who   not   come

    'Someone came, (and) someone didn't come.'

Turning to the scope behaviors of *mzi+wh* strings, they interact with extensional operators in the same way that plain indefinites do. *Mzi+wh* strings may take wide scope over or below a universal quantifier. In (24), a (prominent) wide scope reading is available.[12]

(24) mui-go   hoksaang dou hok-gwo **mzi** bin-zung jyujin

    every-CL student   all   learn-EXP MZI which-CL language

    'Every student has learned some language.'

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

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

Given sufficient contexts, a narrow scope reading of a *mzi*-indefinite is also possible:[13]

(25) Context: Every student in School A is required to take one language course. It must be either German, French or Spanish. The speaker then infers:

    mui-go   hoksaang dou hok-gwo **mzi** bin-zung ngauzau   jyujin

    every-CL student   all   learn-EXP MZI which-CL European language

    'Every student has learned some European language.'

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

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

---

12. The unnatural narrow scope reading implies that the wide scope reading cannot be reduced to a special case of narrow scope reading (see, for discussions, Ruys 1992; Reinhart 1997).

13. The availability of the narrow scope reading seems to be related to whether the alternative set denoted by the *mzi*-indefinite is *d-linked* or not (Pesetsky 1987).

A similar set of examples along this line of discussion concerns negative quantifiers.[14] Mzi+wh strings interact with negative quantifiers such as *mou jan* 'no one' - they can take wide over or narrow scope below the negative quantifiers. (26) favors a wide scope reading, whereas (27) favors a narrow scope reading.[15]

(26) a. Mou   jan   tai-gwo   **mzi** bin-bun   syu

not.have person read-EXP MZI   which-CL book

'No one has read a book (that is unknown to the speaker).'

b. Wide: ∃y[an-unknown-book(y) ∧ ¬∃x[person(x) ∧ read(x,y)]]

c. ??Narrow: ¬∃x[student(x) ∧ ∃y[an-unknown-book(y) ∧ read(x,y)]]

(27) a. Mou   jan   tai-zo   **mzi** bin-bun   syu

not.have person read-PERF MZI   which-CL book

'No one has read a book (that is unknown to the speaker).'

b. ??Wide: ∃y[an-unknown-book(y) ∧ ¬∃x[person(x) ∧ read(x,y)]]

c. Narrow: ¬∃x[student(x) ∧ ∃y[an-unknown-book(y) ∧ read(x,y)]]

Lastly, *mzi+wh* strings also show island-violating scope behaviors, which has been extensively discussed for indefinites (Fodor and Sag 1982, i.a.). For example, they can take scope over a universal quantifier in the matrix clause when occuring within a complex noun phrase island.

(28) mui-go   hoksaang dou tengdou [hokhaau kwaidingjiu hok   **mzi** bin-zung jyujin]

every-CL student   all heard   school   require   learn MZI   which-CL language

ge siusik

GE news

Wide: 'There is some language such that every student heard the news that the school requires (them) to learn it.'

In short, *mzi+wh* strings display canonical scope behaviors of indefinites. This follows without

---

14. I thank an anonymous reviewer for drawing my attention to negative quantifiers.
15. It is unclear to me why the choice of aspect marker may affect the preferred scope reading in the two sentences. I leave this issue open to future research.

further assumption from a choice-functional analysis on *mzi*, which contributes to the existential meaning while allowing scope flexibility (Kratzer 1998; Winter 1997; Reinhart 1997). Before I propose a compositional analysis for this in §4, I first examine the modal/epistemic component associated with *mzi*.

3 The ignorance component as a conventional implicature

In this section, I focus on the interpretation of *mzi*+*wh*-nominals, which I refer to as *mzi*-indefinites. In §3.1, I discuss the basic properties of the ignorance component of *mzi*-indefinites and reveal that there is no anti-singleton requirement (Alonso-Ovalle and Menéndez-Benito 2010). In §3.2, I show that this ignorance component survives various intensional contexts. In §3.3, I argue that the ignorance component of *mzi*-indefinites is not derivable from conversational (e.g., quantity) implicatures. I instead propose a conventional implicature approach in §3.4.

3.1 The ignorance component

An important component in *mzi*-indefinites concerns the epistemic state of the speaker. Intuitively, *mzi* signals that the speaker is ignorant of the referent of the indefinite. Accordingly, *mzi*-indefinites are incompatible with continuations that presume the speaker's knowledge of the referent. For example, a *mzi*-indefinite as in (29a) contrasts with a plain indefinite as in (29b) in that the former cannot be felicitously continued by either sentence in (30), which presumes the speaker's knowledge of the referent.[16]

---

16. The ignorance component is reminiscent of a kind of specificity discussed in Farkas (1995), namely, epistemic specificity (as opposed to scopal specificity). The interpretation of the indefinite in (i-a) is said to be epistemically specific if continued by (i-b), and epistemically non-specific if continued by (i-c).

(i)   a.   A student in Syntax 1 cheated on the exam.

       b.   His name is John.

       c.   We are all trying to figure out who it was.   (Farkas 1995)

In this regard, a *mzi*-indefinite can be said to be epistemically non-specific. However, I stick to the term *ignorance component* to avoid confusions over different types of specificity.

(29)   a.   Aaming tai-zo   **mzi** bin-bun   syu

Aaming read-PERF MZI which-CL book

'Aaming read some book.'   =(2a)

b.   Aaming tai-zo   jat-bun syu

Aaming read-PERF one-CL book

'Aaming read one book.'

(30)   Infelicitous continuations for (29a), but not (29b)

a.   zikhai   Hunglaumung

namely Dream.of.the.red.chamber

'Namely, Dream of the Red Chamber.'

b.   gu-haa hai bin-bun

guess   be   which-CL

'Guess which book it is?'

In some languages, it has been proposed that the ignorance component is incompatible with a singleton domain of quantification. For example, Alonso-Ovalle and Menéndez-Benito (2010) observes that the indefinite marker *algún* 'some' in Spanish requires that the domain of quantification of the indefinite be non-singleton. Consider the Spanish example in (31).

(31) #Juan compró **algún** libro que resultó   ser   el   más   caro   de la libreriía.

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.'

In (31), *más* 'most' in the relative clause forces the domain of *algún* to be a singleton set. Alonso-Ovalle and Menéndez-Benito (2010) take its infelicity as evidence for the anti-singleton requirement imposed by *algún*.[17] However, such an anti-singleton requirement is not observed with *mzi*-indefinites. The Cantonese counterpart of (31) with a *mzi*-indefinite is felicitous.

---

17. This requirement is important in the study of ignorance component sas it appears to crosscut two families of epistemic indefinites. On one hand, Japanese *-ka* (Alonso-Ovalle and Shimoyama 2014) and Tiwa *-pha* (Dawson 2018) are similar to Spanish *algún*. On the other, German *irgendein* (Aloni and Port 2015) does not display the same requirement.

(32) Aaming maai-zo **mzi** bin-bun cyun syudim zeoi gwai ge syu

　　　Aaming buy-PERF MZI which-CL whole bookstore most expensive GE book

　　　(Lit.) 'Aaming bought some book that is the most expensive one in the bookstore.'

One may suggest that 'most' does not always force a singleton domain (e.g., one may say *one of the most expensive books* in English). It is therefore instructive to establish a singleton domain in a different way. For example, it is instructive to see whether *mzi*-indefinites can be used when the speaker is pointing at a particular individual. Consider the scenario below (adopted from Alonso-Ovalle and Menéndez-Benito 2003):

(33) Scenario: L and P are visiting the Math department. They don't know anything about the people working there, and they haven't seen any of them before. They suddenly see an individual, who can be inferred to be a professor, frantically dancing on his desk.

The sentence in (34) shows that *mzi*-indefinites are compatible with such a scenario, hence the lack of an anti-singleton requirement.[18]

(34) P: 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!'

As a side note, it is worth discussing how a singleton domain of an indefinite is compatible with an ignorance component. Specifically, if there is only one potential referent in the domain, can the speaker still faithfully claim ignorance? I suggest the notion of relevance is crucial here. In his discussion of the differences between English indefinites marked with *a* and *some*, Strawson (1974) suggests that

> "the choice of 'some' rather than 'a' embodies what might be called an acknowledgement

---

18. As expected, this is in contrast with *algún*, which is disallowed under the same context:

(i)　#P: ¡Mira! **algún** profesor está bailando encima de la mesa!
　　　　　Look ALGÙN professor is dancing on of the table
　　　'(Pointing at the professor) Look! Some professor is dancing on the table!'

　　　　　　　　　　　　　　　　　　　　　　　(Alonso-Ovalle and Menéndez-Benito 2010)

or recognition of the fact that **the identification supplied, though perhaps the best the speaker can do, might be regarded as inadequate to the circumstances of the case**." (Strawson 1974, p.92, emphases mine)

So it is possible that an identification is *inadequate* to the circumstances (as the speaker perceives), even though there is only one potential referent. I therefore maintain that a singleton domain is compatible with the ignorance component.[19]

3.2   Escaping intensional scope

Another important aspect of the ignorance component associated with *mzi*-indefinites is that it survives different intensional contexts, i.e., their interpretation is independent of intensional operators (i.e., it is scopally specific, in the sense of Farkas (1995)).[20] This property is interesting from a typological point of view, since epistemic indefinites in many other languages show non-uniform interactions with intensional operators and the ignorance component may disappear (see especially discussions in Aloni and Port (2015)). What we observe for *mzi*-indefinites, however, is their lack of interaction with intensional operators.

First, when embedded under attitude verbs, the interpretation of *mzi*-indefinites does not depend on the attitude verbs and has a scopally specific reading. The sentence in (35a) is only compatible with a continuation that suggests a specific doctor that Aafan wants to marry (i.e., (35b), but not (35c)).[21]

---

19. This further suggests the context-sensitive nature of identification. For more discussions on the notion of relevant identification, see Aloni (2001) and Aloni and Port (2015).

20. This may sometimes be regarded as a "wide scope" reading. I reserve the term "wide scope" for interpretations where an indefinite outscopes an extensional operator.

21. This is, for example, in contrast with the Russian epistemic indefinite *WH-nibud'*, which obligatorily takes narrow scope, and gives a scopally non-specific reading (Geist 2008, p.156)

(i) a. Igor' hochet zhenit'sja na kakoj-**nibud**' studentke.   [Russian]
         Igor wants marry   WH-NIBUD'   student
         'Igor wants to marry a [non-specific] student.'

    b. * …On znakom s nej dva goda.   c. …On poka ni s kem ne poznakomilsja.
         'He has known her for two years.'   'He didn't get to know anybody.'
                                    scopally specific   scopally non-specific

(35) a. Aafan soeng tong **mzi** bin-go   jisang gitfan

Aafan want with MZI which-CL doctor marry

'Aafan wants to marry to a [specific] doctor.'

b. ... keoidei sik-zo   loeng-nin

they   know-PERF two-year

'They knew each other for two years.'   scopally specific

c. #... daan keoi jat-go jisang dou m-sik

but   she one-CL doctor all   not-know

'But she does not know any doctor.'   scopally non-specific

Also, *mzi*-indefinites are compatible with epistemic modals, and are likewise interpreted scopally specifically.[22]

(36) Aaming honeng tai-zo   **mzi** bin-bun syu

Aaming possibly read-PERF MZI which-CL book

'There is some book that Aaming may have read.'   scopally specific

When embedded under deontic modals, a *mzi*-indefinite is still interpreted as a scopally specific indefinite, with an ignorance component.[23]

---

22. In Czech, epistemic indefinites (marked with *jakémsi*) display what Šimík (2014) calls *epistemic clash*, where they cannot be embedded under epistemic modals, as opposed to the one marked with *nějakém* 'some'.

(i)   Určitě/ možná spí   na { nějakém/ ***jakémsi** } gauči.   [Czech]
surely maybe sleep.3sg on   some   some.EI   couch
'Surely/Maybe he's sleeping on some couch.'

23. Interestingly, the German *irgendein*-marked indefinite can have a free choice reading, and the ignorance component ceases to exist (Kratzer and Shimoyama 2002; Aloni and Port 2015). The following example is from Kratzer and Shimoyama (2002, p.10-11). Note that the specific reading is also available.

(i)   Mary musste **irgendeinen** Mann heiraten.   [German]
Mary had-to IRGEND-one   man   marry
a. 'Mary had to marry a man, any man was a permitted marriage option for her.'   scopally specific
b. 'There was some man Mary had to marry, the speaker doesn't know or care who it was.'   free choice

(37) 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

The same goes for conditional sentences. When *mzi*-indefinites appear in the conditional antecedent, the ignorance component persists. The sentence in (38) conveys a reading that there is a particular book unknown to the speaker, such that Aaming will come after finish reading it.[24]

(38) jyugwo Aaming tai-jyun   **mzi** bin-bun   syu,   keoi zau   wui lai

if   Aaming read-finish MZI which-CL book, he   then will come

'There is some book such that if Aaming finishes reading it, he will come.'

Interestingly, the ignorance component in *mzi*-indefinites also survives/"outscopes" the illocutionary force of a sentence, patterning with what Dawson (2020) observes for epistemic indefinites in Tiwa. For example, in (39), the sentence conveys a pragmatically marked but not impossible meaning: the speaker knows that Aaming wanted to see a particular person, but they can't remember whether it was Ms Lee or Ms Wong. What the speaker wants to know by uttering (39) is whether Aaming met that particular person (still unknown to the speaker).

(39) The ignorance component persists in *interrogative* clauses

Aaming jau-mou   gin **mzi** Lei lousi ding Wong lousi?

Aaming have-not.have meet MZI Ms Lee   or   Ms Wong

'Did Aaming see Ms Lee or Ms Wong?"

Similarly, in (40), the speaker knows that the addressee needs to go and talk with a particular teacher, but the speaker can't remember whether that teacher is Ms Lee or Ms Wong. The speaker instructs the addressee to go and talk with whoever it is they need to.[25]

---

24. I thank an anonymous reviewer for raising this issue.

25. For further discussions on these examples and implications on imperatives, see Dawson (2020).

(40)   The ignorance component persists in *imperative* clauses

   heoi gin   **mzi** Lei lousi ding Wong lousi laa1

   go   meet MZI Ms Lee   or   Ms Wong   SFP

   'Go and see Ms Lee or Ms Wong!'

The examples above demonstrate that *mzi*-indefinites are always scopally specific and the ignorance component persists as well. This is reminiscent of certain types of indefinites as in St'a̓ t'imcets (Matthewson 1999) and Tiwa (*-khi*, Dawson 2018), where such elements take obligatory wide scope. One difference, however, is that *mzi*-indefinites can in fact take narrow quantificational scope, as illustrated in (25). This also suggests that the scope behaviors of *mzi*-indefinites are not uniform with regard to extensional and intensional quantifiers.

Two remarks concerning the notion of specificity are in order. First, in the above examples, it may be tempting to attribute the scopal specificity to *bin* 'which', due to its property of being *discourse-linked* (Pesetsky 1987). The example below however shows that the scopally specific reading remains if the *wh*-expression is replaced by *matje* 'what'. The same applies to a colloquial version of 'who', namely, *matseoi*.[26] I therefore conclude that it is *mzi* that is responsible for scopal specificity.

(41)   Aafan soeng tong {**mzi** matje jisang/ **mzi** matseoi} gitfan

   Aafan want with MZI   what doctor/ MZI who   marry

   'There is some kind of doctor/ someone that Aafan wants to marry to.'   scopally specific

Second, if the ignorance component is construed as epistemic non-specificity (following Farkas (1995), see also fn.16), *mzi*-indefinites constitute further evidence for the claim that two kinds of specificity may be involved which are independent of each other: an indefinite can be epistemically non-specific but scopally specific.[27]

---

26. I switch to a colloquial form because, morphology-wise, the more common version *bin-go* 'who (lit. which-CL) also contains *bin-* 'which'.

27. Relatedly, Farkas (1995) raises the issue of whether the opposite exists, i.e. whether there is a kind of indefinite that is epistemically specific but scopally non-specific. The question is relevant to the interpretation of the following sentences, which would not be addressed here.

3.3 Not a conversational implicature

A well-received approach to the ignorance component in epistemic indefinites is to treat this as a conversational implicature derived via Gricean reasoning. The idea at least goes back to the discussion of (42) in Grice (1975). It is suggested that the ignorance component indicates that, on one hand, the speaker is aware of being less informative than required (infringement of the maxim of Quantity). On the other hand, however, "to be more informative would be to say something that infringed the maxim of Quality," (p.51) since the speaker may make a false claim. As a result, in the exchange in (42), even though a better answer would be, for example, Nice or Eze, with the use of *somewhere*, the speaker "implicates that he does not know in which town C lives" (p.52).

(42)   A: Where does C live?

   B: Somewhere in the South of France.

   Similar approaches have been adopted and developed in Alonso-Ovalle and Menéndez-Benito (2010), Chierchia (2013), Chierchia and Liao (2015), Alonso-Ovalle and Shimoyama (2014), and Dawson (2018), among others. In what follows, I argue that the ignorance component associated with *mzi*-indefinites should not be treated as such, for it does not display the signature properties of a conversational implicature.

   First of all, the ignorance component in a *mzi*-indefinite cannot be canceled or reinforced.[28] The continuation that explicitly cancels the ignorance gives rise to an air of incoherence.

(43)   <u>Non-cancellability</u>

   #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).'

   In addition, reinforcing the ignorance component in a subsequent sentence results in redundancy

---

(i)   a.   John believes that a unicorn has destroyed his flowerbeds.

   b.   Cob believes that a witch has blighted his mare and Nob believes that she has destroyed his crops.

28. The opposite is observed with Spanish and Japanesee epistemic indefinites (see, for example, Alonso-Ovalle and Menéndez-Benito 2010; Alonso-Ovalle and Shimoyama 2014, respectively)

(Horn 1972; Sadock 1978).

(44)   Non-reinforceability

#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).'

Note that if we replace the *mzi*-indefinite with a plain indefinite such as *jat-bun syu* 'one-CL book' in (43) and (44), both sentences become felicitous.

Second, it is often suggested that a conversational implicature tends to disappear in downward entailing environments (Gazdar 1979; Horn 2001; Chierchia 2004). However, the ignorance component of *mzi*-indefinites persists when embedded under negation. Note that the *mzi*-indefinite scopes over the negation and does not result in an NPI reading.

(45)   Aaming mou   tai-gwo   **mzi** matje syu

Aaming not.have read-EXP MZI what book

'There is some kind of books that Aaming hasn't read.'

($\neq$ 'Aaming hasn't read any kind of books.')

Similarly, the ignorance component remains when embedded under the attitude verb *waaiji* 'doubt'.[29]

(46)   Aafan waaiji Aaming jau-mou   tai **mzi** bin-bun   syu

Aafan doubt Aaming have-not.have read MZI which-CL book

'There is some book such that Aafan doubts whether Aaming read it.' ($\neq$ 'Aafan doubts whether Aaming read any book.')

Third, it has been suggested that the ignorance component that is derived from a conversational implicature does not have to be speaker-oriented. It can be shifted to some other epistemic agent. This is the case for Spanish, where we can see that the ignorance component is not anchored to the speaker, but the subject of the attitude verb (= Juan):

---

29. This is consistent with the claim that *mzi*-indefinites are scopally specific, discussed in §3.2.

(47) Juan sabe   que María se casó   con **algún** estudiante del   departamento.

Juan know:3s that María SE marry:past3s with ALGÚN student   of.the department

  Él no sabe   con quién, ¡pero yo sí!

  He not know:3s with whom, but   I   do

'Juan knows that María married a student in the department. He doesn't know who, but I do!'

<div align="right">(Alonso-Ovalle and Menéndez-Benito 2010, p.13)</div>

By way of contrast, the ignorance component of *mzi*-indefinites cannot be shifted. In (48), a *mzi*-indefinite is embedded under an attitude verb. If the component could be shifted to, e.g., the subject *Aaming*, the sentence would allow a continuation that comes with the identification of the doctor that is provided by the speaker. However, this is not the case.

(48) Aaming gokdak keoi gindou **mzi** bin-go   jisaang (#zikhai Wong jisaang)

Aaming think   he   saw   MZI which-CL doctor namely Wong doctor

'Aaming thinks that he saw some doctor (namely Dr. Wong).'

The infelicity of the *namely*-continuation suggests that the ignorance component of a *mzi*-indefinite is consistently oriented to the speaker, but not the subject (another potential epistemic agent).

Similarly, while an 'according to...' phrase introduces a salient epistemic agent, the ignorance component is still oriented to the speaker. In (49), it is the speaker, but not the government, who cannot identify the referent mentioned by the government.

(49) gangeoi   zingfu   ge gongfaat, Hoenggong gingzai hou-gwo **mzi** bin-go

according.to government GE saying   Hong.Kong economy good-than MZI which-CL

  singsi

  city

'According the government's saying, Hong Kong's economy is better than some city.'

So a *mzi*-indefinite always conveys the speakers' ignorance. This is different from what is observed for an ignorance component that is characterized as a conversational implicature.

In short, the kinds of evidence that have elsewhere have been given for a conversational/quantity

implicature approach do not apply to *mzi*-indefinites. I conclude that the ignorance component of *mzi*-indefinites is not derivable from a conversational implicature.[30] [31]

3.4   Conventional implicature

In the previous subsections, we have observed that the ignorance component of *mzi*-indefinites displays the following properties:

(50)   Properties of the ignorance component of *mzi*-indefinites

    a.   it is compatible with a singleton domain of quantification (i.e., it can refer to a specific unknown);

    b.   it survives various intensional contexts;

    c.   it cannot be canceled or reinforced;

    d.   it does not disappear in downward entailing environments;

    e.   it is always oriented to the speaker.

Based on these observations, I propose that the ignorance component of *mzi*-indefinites is best regarded as a *conventional implicature*. I adopt the definition of conventional implicatures in Potts (2005, 2007), which basically follows the suggestions in Grice (1975).

(51)   Conventional implicatures (CIs; Potts 2005, p.11)

    a.   CIs are part of the conventional meaning of words.

    b.   CIs are commitments, and thus give rise to entailments.

    c.   These commitments are made by *the speaker of the utterance* 'by virtue of the meaning of'

---

30. The absence of an anti-singleton requirement in *mzi*-indefinites discussed in §3.1 provides a further argument against a conversational/quantity implicature approach. This is because if the quantification domain is a singleton set, it would not violate the maxim of Quantity (see, for discussions along this line of reasoning, Alonso-Ovalle and Menéndez-Benito 2010).

31. Another variant of the conversational implicature approach suggests that the ignorance component may be a *manner* implicature, which is attributable to lexical competition with another expression (cf. lexical blocking, McCawley 1978), as in Russian (Geist 2008) and Tiwa (Dawson 2018). In view of the absence of obvious competitors to *mzi*, I do not consider this possibility.

the words he chooses. (emphasis original)

d. CIs are logically and compositionally independent of what is 'said (in the favored sense)', i.e. independent of the at-issue entailments.

The ignorance component of *mzi*-indefinites is obviously encoded by *mzi* given its morphological makeup. We have also seen that it cannot be canceled - it is entailed by *mzi*. Also, the ignorance component is committed by the speaker, i.e., it is always oriented to the speaker. These observations satisfy both (51a), (51b), and (51c).

As for (51d), to see whether the ignorance component of *mzi*-indefinites is logically independent of the *at-issue* content of the sentence, I adopt the 'yes, but...'-test to illustrate this point (Karttunen and Peters 1979; Potts 2005). In the exchange in (52), after hearing (52a), the hearer can follow up by first agreeing on the at-issue existential claim with *hai aa3* 'yes' (i.e., Aaming read some book), but go on to disagree with the ignorance component with the *batgwo* 'but'-clause (i.e., the speaker does not know which book it is'). Disagreeing with the ignorance component after saying *hai* 'yes' does not render B an incoherent interlocutor.

(52) a. A: Aaming tai-zo **mzi** bin-bun syu =(2a)

Aaming read-PERF MZI which-CL book

A: 'Aaming read some book.'

b. B1: hai aa3, batgwo nei zi hai bin-bun gaak3

yes SFP but you know be which-CL SFP

B1: 'Yes, but you know which book it is.'

Alternatively, the hearer may suspend the discussion of the *at-issue* content and reinvoke the ignorance component as an item for discussion. S/he may disrupt the flow of the conversation with, for example, a "hey, wait a minute" sentence (Shanon 1976; Fintel 2004).

(53) B2: dang zan sin, nei zi hai bin-bun gaak3

wait a.moment first, you know be which-CL SFP

B2: 'Wait a minute. You know which book it is.'

As such, the ignorance component of *mzi*-indefinites is logically independent of the at-issue content. Together with the arguments against its status as a conversational implicature, I conclude that the ignorance component of *mzi*-indefinites should be regarded as a conventional implicature.[32]

As a final remark, a conventional implicature approach is not the only non-Gricean account in the literature of epistemic indefinites. For example, Jayez and Tovena (2006) and Giannakidou and Quer (2013) suggests that the use of epistemic markers (e.g., French *un quelconque* and Greek *-dhipote*) is constrained by a felicity condition, namely, the *Referential Vagueness Condition*. It states that with the use of the referential vagueness marker, the speaker suggests that s/he does not intend to refer to exactly one individual. But this is not the case for *mzi*, since we have seen that *mzi*-indefinites can be used felicitously even though there is only one potential referent (see §3.1).

4   A compositional analysis

This section sketches a (semi-)formal compositional analysis of *mzi*, based on the findings in both §2 and §3. I first make explicit a few assumptions and then go on to implement the idea that *mzi* is a choice function that triggers a conventional implicature relating to the speaker's ignorance. Then I turn to a complication on what I call *distributed ignorance*.

4.1   Assumptions

First, given our discussion in §2.1, I assume the internal structure of a *mzi*-phrase as in (54), which indicates that *mzi* is syntactically a modifier, or an adjunct-like element. I use αP to indicate the category flexibility of *mzi* (recall that it can combine with elements of different syntactic categories). I use the label *WH* to cover both *wh*-expressions and disjunctive phrases.

(54)   <u>The internal structure of a *mzi*-indefinite</u>

$[_{\alpha P}$ *mzi* $[_{\alpha P}$ *WH* $]]$

---

32. It is also possible to suggest that the ignorance component of *mzi*-indefinites is in fact a *presupposition*. I do not make further distinction between presupposition and conventional implicature (but see discussions in Potts (2005)).

Second, I assume an alternative-based approach to both *wh*-expressions and disjunctive phrases, following Hamblin ([1973](#)), Kratzer and Shimoyama ([2002](#)), Beck ([2006](#)), Cable ([2010](#)), Simons ([2005](#)), and Alonso-Ovalle ([2006](#)). Accordingly, the ordinary semantic values of both kinds of expressions are undefined, but they invoke sets of alternatives for their evaluation. For examples, *bin-bun syu* 'which book' invokes a set whose members are books, whereas *syu ding bouzi* 'book or newspaper' a set whose members include books and newspapers. Their semantics can be given as follows:

(55)   a. $[\![$*bin-bun syu*$]\!]$ = {x: book(x)}

    b. $[\![$*syu ding bouzi*$]\!]$ = { book, newspaper }

Lastly, I assume a multi-dimensional semantic framework (Karttunen and Peters [1979](#); Potts [2005](#)), where single sentences can express more than one, non-conjoined proposition. Substantially, the meaning of a sentence can be two-dimensional, where the first dimension concerns the *at-issue content* and the second dimension the (non-at-issue) conventionally implicated content.[33] I follow the formalization of conventional implicature in Potts ([2005](#)) (what he calls $\mathcal{L}_{CI}$, a CI logic). The distinction between the two types of meaning is implemented via the set of semantics types, given in (56).

(56)   <u>The definition of the set of types for $\mathcal{L}_{CI}$</u>

    a. e^a, t^a, and s^a are basic at-issue types.

    b. e^c, t^c, and s^c are basic CI types.

    c.   If σ and τ are at-issue types, then <σ, τ> is an at-issue type.

    d.   If σ is an at-issue type and τ is a CI type, then <σ, τ> is a CI type.

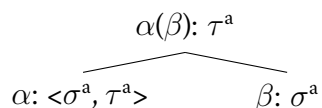   e.   The full set of types is the union of the at-issue and CI types.

With regard to semantic composition, there are two applications in $\mathcal{L}_{CI}$. *At-issue* application, as given in (57a), is a version of the rule for functional application of sisters of Heim and Kratzer ([1998](#)). On the other hand, CI application, as given in (57b), allows a functional CI term (= α) to take an at-issue term (= β) and yield part of the value of their mother as CI, while passing on the unmodified

---

33. For formal implementation of this idea, see Karttunen and Peters ([1979](#)) and Potts ([2005](#)).
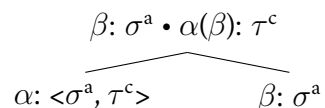
value of the at-issue term. This yields two dimensions of meanings, separated by the bullet • in (57b).

(57)   Application in $\mathcal{L}_{CI}$ (Potts 2005)

    a. At-issue application

$$\alpha(\beta):\tau^a$$
$$\alpha:<\sigma^a,\tau^a> \qquad \beta:\sigma^a$$

    b. CI application

$$\beta:\sigma^a \bullet \alpha(\beta):\tau^c$$
$$\alpha:<\sigma^a,\tau^c> \qquad \beta:\sigma^a$$

4.2   Proposal and implementation

4.2.1   *Mzi* as a choice function variable and a CI trigger

Based on the observations in §2, I suggest that *mzi* is both an overt realization of a choice function variable (Kratzer 1998; Reinhart 1997; Winter 1997), as well as a CI trigger of the ignorance component (Potts 2005). On one hand, the choice function is conveyed as at-issue content. A choice function is a function that takes a (non-empty) set denoted by descriptive content as its argument and returns a unique member of that set, formally defined as in (58).[34]

(58)   Let f be a choice function (variable),

$$f =_{DEF} \forall X[\, X \neq \varnothing \to f(X) \in X]$$

Given the category flexibility of *mzi* discussed in §2.1, I suggest that the semantic type of X may range over entities, properties or propositions, rather than <e,t>-type or intensionalized <s,<e,t>>-type as originally proposed. It follows the existential meaning of *mzi*-indefinites constitutes the at-issue content.

On the other hand, given that the ignorance component is best regarded as a (non-at-issue) conventional implicature, I suggest that *mzi* is a CI trigger at the same time - it is a functional CI term that take an at-issue term as its argument and return a proposition (as CI). Roughly, the proposition conveys that the speaker doesn't know the referent chosen by the choice function.

To implement this idea in the multi-dimensional framework in $\mathcal{L}_{CI}$, let us consider the example in (2a), again,

---

34. The definition is taken from Geurts (2000), which is based on Reinhart (1997).

(59) Aaming tai-zo   [**mzi** bin-bun   syu]   = (2a)

Aaming read-PERF MZI   which-CL book

⤳ 'Aaming read some book (I don't know which).'

I suggest that following lexical semantics of *mzi* as given in (60). Notably, *mzi* is both an at-issue type and a CI type, both taking the same type as its arguments.

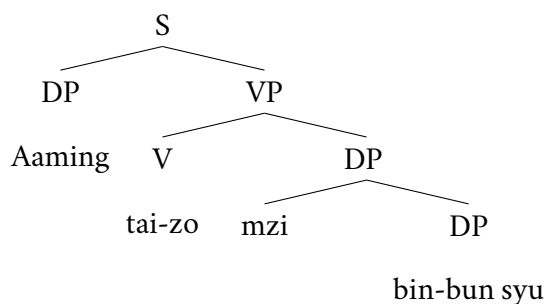(60) A multi-dimensional semantics of *mzi*

    a. *At-issue content*:

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

    b. *Conventional implicature*:

        $[\![mzi_i]\!]^g = \lambda P_{<e,t>}.\ \text{not-know}(Speaker, g(i)(P))$ $<<e^a, t^a>, t^c>$ (a CI type)

The meaning of (59) can be captured by the semantic composition in (61b), based on the structure in (61a). The crucial semantic contribution is put in boxes, which returns to both an entity (of at-issue type) and a proposition (of CI type), after applying at-issue application and CI application to its argument (i.e., the *wh*-expression). Ultimately, the sentence conveys the at-issue content of *Aaming read some book* and the non-at-issue content of *the speaker doesn't know this book*.

(61) a. Syntactic structure of (59)

```
                      S
            ┌─────────┴─────────┐
           DP                  VP
         Aaming        ┌────────┴────────┐
                       V                DP
                     tai-zo      ┌───────┴───────┐
                               mzi              DP

                                          bin-bun syu
```

    b. Semantic composition of (59)

READ(Aaming, g(i)({x: book(x)})) : t^a

Aaming: e^a λ**y. READ(y, g(i)({x: book(x)}))**: $<e^a, t^a>$

tai-zo: $<e^a, <e^a, t^a>>$ **g(i)({x: book(x)})**: e^a ·

NOT-KNOW(*Speaker*, g(i)({x: book(x)})): t^c

mzi: $<<e^a, t^a>, e^a> · <<e^a, t^a>, t^c>$ **{x: book(x) }**: $<e^a, t^a>$
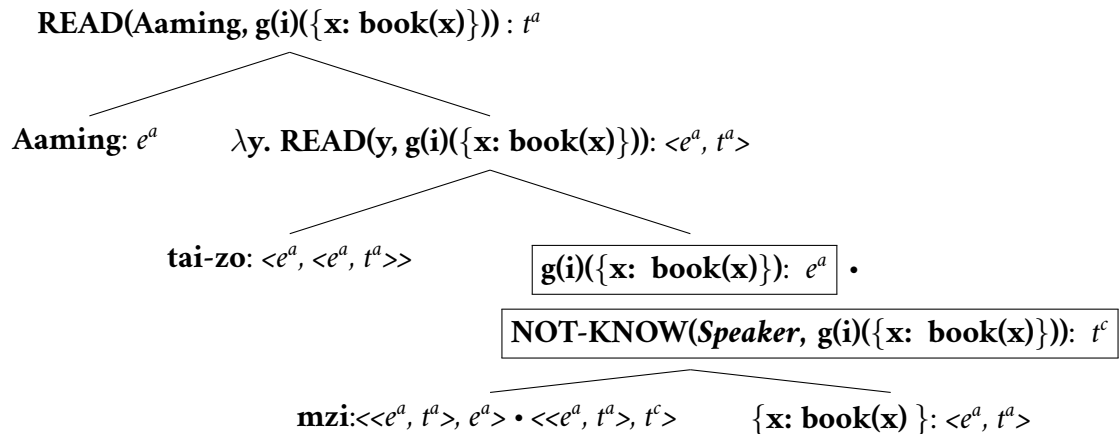
Following Kratzer (1998), I assume that the choice function is contextually bound, instead of being existentially bound at different levels of the sentence (Winter 1997; Reinhart 1997). In other words, the context of use would determine a choice function as the denotation of *mzi*. The whole string *mzi bin-bun shu*, then, denotes the book that the choice function picks from the set of all (relevant) books. This implementation of choice function entails a wide-scope reading of *mzi*-indefinites, and readily capture the (exceptional) wide scope properties of a *mzi*-indefinite discussed in §2.2.3 and in §3.2.

4.2.2 The apparent narrow scope and parameterized choice functions

It should be remarked that *mzi*-indefinites do not always take the widest scope in a sentence. We have already discussed examples in §2.2.3 and witnessed that in some cases *mzi*-indefinites can have a narrow scope reading with regard to universal quantifiers (at least apparently). The relevant example in (25) and the target reading are repeated below.

(62) a. mui-go   hoksaang   dou   hok-gwo   [**mzi** bin-zung   auzau   jyujin]   =(25)

       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-<u>unknown</u>-European-language(y) ∧ learned(x,y)]]

Here, I follow Kratzer (1998) again in assuming that the (apparent) narrow scope reading in cases like (62) can be derived via a combination of choice functions and implicit arguments (i.e., *paramater-*

ized choice function). Concretely, the choice function is associated with an implicit argument, which can receive a referential or a bound variable interpretation.

(63) Parameterized choice functions (Kratzer 1998)

    a.   The choice function is a pronominal element f_x;

    b. f is an unbound variable subject to a contextually determined assignment;

    c.   its pronominal argument x is contextually determined or bound by a wider scope quantifier.

As such, the narrow scope reading is indeed an apparent one. Instead, it is a wide scope functional reading, where the choice function takes wide scope, but the pronominal argument on the choice function is bound by a wide scope (universal) quantifier. In our case in (62), the pronominal argument on the choice function is bound by the universal quantifier *mui-go hoksaang* "every student".

(64) The at-issue meaning of (62)

    $\forall x[\text{student}(x) \rightarrow \text{learned}(x, f_x\{\text{Spanish, German, ... }\})]$

One support for adopting the Kratzer's approach to choice functions choice function comes from the lack of intermediate scope. Since Kratzerian style's choice function variable is left unbound, it predicts the absence of (genuine) intermediate reading. This is borne out in the following case of *mzi*-indefinites. In (65), while it is possible to have a wide scope or the (apparent) narrow scope reading, the intermediate scope reading is unavailable.

(65) mui-go   doujin   dou jiucing-gwo daiboufan taiming-zo **mzi** bin-tou   dinjing ge

     every-CL director all   invite-EXP   most   nominate   MZI which-CL film   GE

       pingleongaa

       critic

    a. Wide: There is some film such that every director invited most film critics who nominated it.

    b. ??Intermediate: For every director$_i$, there is some film such that he$_i$ invited most critics who have nominated it.

    c. Narrow: For every director$_i$, he$_i$ invited most critics who nominated some film.

This is consistent with the findings in Mandarin, which also lacks a genuine intermediate scope reading, as discussed in Kim (2003).[35]

4.3 Distributed ignorance

4.3.1 A complication

There is an intriguing complication relating to the ignorance component when *mzi*-indefinites are interpreted under a functional wide scope reading. There are cases where the ignorance component conveyed by a *mzi*-indefinite is distributed to each member in the domain of the wide scope universal quantifier. To establish a concrete case of how the ignorance component can be distributed, consider (66) below. Assuming that every Japanese song is only rearranged once, the sentence has a preferred "narrow" scope reading of the *mzi*-indefinite (= a functional wide scope reading).

(66) a. mui-sau 80nindoi coetmeng ge go dou hai goipin zi **mzi** bin-sau

every-CL eighties famous GE song all be rearrange from MZI which-CL

jatman-go

Japanese-song

'Every famous song in the eighties is rearranged from some Japanese song.'

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

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

Importantly, this also illustrates a case of what I call *distributed ignorance*: the speaker is ignorant on all the pairs of famous songs and Japanese songs, i.e., the speaker doesn't know, for each famous song, which Japanese song it is re-arranged from. In other words, for each famous song that co-varies with an unknown Japanese song, the speaker cannot identify that Japanese song.

Before I turn to potential solutions to this issue, this property of the ignorance component is, as far as I know, not observed in epistemic indefinites in other languages. It should, however, be noted that Alonso-Ovalle and Shimoyama (2014) examine the interaction between epistemic indefinites in

---

35. *Mzi*-indefinites differ from English indefinites in this regard, where genuine wide scope is possible, see discussion in Endriss (2009, p.159-160).

Japanese (in the form of *WH-ka*) and universal quantifiers, but they observe that the ignorance component *disappears* when *WH-ka* is interpreted narrowly. Consider the following scenario (68a), adopted from Alonso-Ovalle and Shimoyama (2014):

(67)   Every Friday afternoon, students and professors at the Philosophy Department practice tango. Right now, every professor is dancing with a student. J is seeing quite clearly the scene. She knows perfectly well who is a student and who is a professor in that department. L calls J over the phone.

The scenario is set up such that J has perfect information on all the professor-student pairs. Alonso-Ovalle and Shimoyama (2014) observe that the use of *WH-ka* is still felicitous as in (68a). Furthermore, the follow-up questions by L in (68b), which presupposes J's knowledge of the professor-student pair, is also appropriate.

(68)   Context: (67)

a.   J: Dono kyooju-mo   dare-**ka** gakusee-to   odotteru.

 which professor-мо who-ка   student-with is.dancing

'Every professor is dancing with some student.'

b.   L: Dare-ga   dare-to   odotteru   no?

 who-ɴᴏᴍ who-with is.dancing Q

'Who is dancing with who?'   (p.13)

The acceptability of (68a) indicates the absence of the ignorance component, since, if it existed, (68a) would constitute a contradiction and the follow-up in (68b) would render L an uncooperative speaker (as ignorance is already claimed by J). In contrast, under the same scenario, the Cantonese counterpart with a *mzi*-indefinite suggests the opposite, where the ignorance component of a *mzi*-indefinite does not disappear when interpreted under a universal quantifier. It thus results in infelicity in (69).

(69) Context: (67)

  #J: mui-go gaausau dou tung-gan **mzi** bin-go hoksaang tiumou

    every-CL professor all with-PROG MZI which-CL dance

  'Every professor is dancing with some student.'

The difference between epistemic indefinites in Cantonese and Japanese lends further support to the different characterization of the ignorance component in the two language. As proposed, *mzi*-indefinites convey a conventional implicature on speaker's ignorance and thus persists under a universal quantifier, whereas the ignorance component conveyed by *WH-ka* is arguably a conversational implicature (Alonso-Ovalle and Shimoyama 2014) and can be canceled when interpreted below a universal quantifier. The remaining issue, however, is how this distributed ignorance can be derived under the current proposal.

4.3.2 Towards a solution

The case of distributed ignorance does not immediately follow from the aforementioned Potts' framework on conventional implicature - it is unclear how the universal quantifier can bind into the ignorance component that is in a different dimension of semantics. Indeed, Potts observes that binding into a variable into supplements such as *As*-parentheticals and appositive relative clauses in English is impossible.

(70) No binding into (some) appositives (Potts 2005, p.80, 82)

  a. *No **reporter**$_i$ believes that, as **he**$_i$ wrote, Ames is a spy.

  b. *No **reporter**$_i$ believes that Ames is a spy, which **he**$_i$ wrote in his column.

Potts argues that the unacceptability of (70) follows from the fact that the variable in the supplements are left unbound (in the non-at-issue dimension) and thus interpreted as a free pronoun (see discussions Potts 2005, p.79-82).

It is tempting to suggest that the same reasoning need not apply to the case of *mzi*, since the choice function denoted by *mzi* is left unbound in the first place, and since the choice functions in different

dimensions are to be retrieved their value from the same salient context. However, with the introduction of parameterized choice functions, the implicit argument would require a binder in the same dimension, running into a similar issue as Potts suggests for English supplements.

Here, it should be noted that binding into supplements in English is not strictly ruled out in all cases - there appear to be cases of successful binding. Consider a case reported in Ott (2016) below.

(71)   Binding into (some other) appositives (p.29)

Every **inmate**$_i$ talks to one person, (probably) **his**$_i$ mother, once a week.

Instead of suggesting that the universal quantifiers directly bind the pronoun in the appositives, Ott (2016) makes an analogy between appositives in (71) and the fragment answer in (72), suggesting that both of them may involve an elided structure, and the pronominal is actually bound not by the binder in the host clause, but by an elided binder in a separate clause that is entirely parallel to host clause.

(72)   a.   A: Who does [**every inmate**]$_i$ talk to once a week?

       b.   B: (Probably) **His**$_i$ mother.

A similar analogy can be made with *mzi* in Cantonese as well. Consider a case in (73) where *mzi* is used in a fragment answer. (73) conveys a meaning such that the speaker thinks that every famous song is rearranged from a (different) Japanese song, but expresses his/her ignorance over each of these songs. In other words, the ignorance component is distributed in a similar way to the example in (66).

(73)   a.   A: **mui-sau 80nindoi coetmeng ge go**$_i$ dou hai goipin   zi   bin-sau   go

            every-CL eighties   famous   GE song all   be   rearrange from which-CL song

            aa3?

            SFP

            Lit.: 'Every famous song in the eighties is rearranged from which song?'

       b.   B: **mzi bin-sau   jatman-go**$_i$ gwaa3.

            MZI which-CL Japanese-song SFP

            'Perhaps some Japanese song.'

Accordingly, I suggest that the case of distributed ignorance exhibited with *mzi* can receive a parallel treatment as the kind of appositives discussed in Ott ([2016](#)), which similarly involve an elided structure that parallel to the host clause. Binding between a quantifier and the implicit argument in the choice function may occur at the at-issue dimension before the CI functional *mzi* applies, resulting in distributed ignorance as shown in cases like ([66](#)) and ([73](#)).

5 Notes on the grammaticalization path of *mzi*

5.1 Different usages of the string *m-zi*

Throughout this paper, I have argued that, from a synchronic perspective, the non-verbal *mzi* is substantially different from the verbal counterpart. However, this is not to say that the two usages of *mzi*/*m-zi* are unrelated diachronically. Here I discuss a potential grammaticalization path from the predicative *m-zi* to the choice functional *mzi*, specifically focusing on the change of the selectional requirement of *m-zi*.

First of all, in addition to the two usages of *mzi*/*m-zi* already discussed, it is crucial to note that Yap and Chor ([2014](#)) further observe that *m-zi* can be used as a (negative) attitudinal marker, which conveys the speaker's indifference to the content of the interrogative clause. To see an example,

(74) Aaming **m-zi**   tai-zo   bin-bun syu

      Aaming not-know read-PERF which-CL book

      '(I) don't know which book Aaming read.' $\neq$ 'Aaming does not know which book (he) read.'

Notice that the surface subject is *not* thematically related to *m-zi* (but to the embedded verb) and the (unexpressed) epistemic agent is by default taken to be the speaker (i.e., it conveys the speaker's ignorance or indifference). This usage of *m-zi* retains the clause-taking ability of the verbal *m-zi* but its external argument is suppressed/dropped (or becomes unnecessary). I tentatively call such usage of *m-zi* a *raising* one.[36] Interestingly, ([76](#)) indicates that the clausal complement of this *m-zi* has to be an

---

36. This raising usage in ([74](#)) is not unique to *m-zi*, but is observed with a wider set of attitude verbs, examples of which have been reported in Lee and Yip ([2021](#)).

interrogative one, as opposed to its attitude verb usage (see (1)).

(76) *Aaming **m-zi** tai-zo ni-bun syu

Aaming not-know read-PERF this-CL book

Int.: '(I) don't know Aaming read this book.'

The different usages and the selectional requirement of *m-zi* are summarized as follows:[37]

	External argument	Internal argument	Examples
(a) As an attitude verb	present	declarative or interrogative clauses	(1)
(b) As a "raising" verb	absent	interrogative clauses	(74)
(c) As a choice function	absent	interrogative clauses or phrases	(2), (16), etc.

Table 2: Different usages of the string *m-zi*

A possible way to interpret the relations between these usages of *m-zi*/*mzi* is that *m-zi* originates as a (two-place) attitude verb in (a). Then, the selectional requirement on the external argument becomes optional and *m-zi* acquires the usage in (b). Additionally, the potential size of internal arguments extends from clauses to phrases, while both are still required to be interrogative (i.e., alternative-invoking expressions (of different syntactic categories)), and this gives rise to the choice functional usage in (c).[38] [39] Notably, the change of the syntactic distribution of *m-zi*/*mzi* is correlated with the development of its semantic properties. Precisely, the lexical meaning/at-issue content of the verbal *m-zi* (i.e., not know) in (a) and (b) becomes a non-at-issue content (i.e., a conventional implicature, as argued) in the choice functional usage in (c). The development is reminiscent of a common grammaticalization process, namely, *de-categorialization* (Hopper 1991, p.22), where a clause taking verb loses

(75) Aaming **gamgok**/ **gugai** tai-zo jat-bun syu
Aaming feel.like/ guess read-PERF one-CL book
'(I) feel/guess Aaming read one book.' ≠ 'Aaming feels like/ guesses that (he) read one book.'

This indicates that *m-zi* as a raising verb should be discussed together with these other cases. I do not pursue an analysis in this paper.

37. *M-zi* can also take NP complements. I set aside this usage.

38. The requirement of interrogative clauses in (b) remains mysterious. It may be related to evidentiality; see Lee and Yip (2021) for discussions.

39. This explains why the elements associated with *mzi* must be interrogative, as suggested in the syntax-semantic requirement of *mzi* as discussed in (18). The requirement is inherited from the usage in (b).

its morpho-syntactic properties and then develops as a functional category.[40] [41]

5.2   Morphological fusion of verbs and *wh*-expressions

In support of the proposed grammaticalization path, diachronic data show that it is commonplace for *wh*-expressions to develop into indefinites by fusing with verbs. In the cross-linguistic investigation of indefinite pronouns, Haspelmath (1997) reports that one of the major sources for indefiniteness markers arises from the combination of elements with the clause 'I don't know', especially in European languages. The following are some examples taken from Haspelmath (1997, p.131):

(77)   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'

I suggest that it is plausible that *mzi* is fused with the *wh*-expressions in a similar way. This suggestion is further supported by the observation that *mzi* indeed requires strict adjacency with its associate. For example, in (78), number-classifier expressions and possessors cannot be inserted between *mzi* and the *wh*-expressions.[42] These sentences are otherwise acceptable if *mzi* follows these expressions and appear adjacent to the *wh*-expressions.

---

40. A case in point is the English epistemic adverb *maybe*.

41. As pointed out by an anonymous reviewer, the Mandarin counterpart of *mzi*, namely *bu-zhi*, differs from Cantonese *mzi* in multiple aspects on the non-predicative usage under discussion. The general observation is that *mzi* enjoys a wider distribution than *bu-zhi*, at least in its interaction with negation, conditionals, and its compatibility with nominal *wh*-expressions. This calls for an independent analysis on the Mandarin *bu-zhi*, which, however, goes beyond the scope of this paper. While I believe that the two lexical items are related, my speculation is that the differences between the two is due to the degree of grammaticalization. I suggest that *mzi* has a greater degree of grammaticalization than *bu-zhi*, given its wider distribution. Furthermore, I also speculate that the reason why *mzi* is grammaticalized to a greater extent than *bu-zhi* is correlated with the more restricted distribution of the existential reading of *wh*-expressions in Cantonese, as discussed in the beginning of the paper in (4) and (5). A more comprehensive study on the variation of epistemic indefinites in Chinese languages is much desired.

42. There is, indeed, one case where *mzi* is separated from its associate, i.e., when there is a plural generic classifier *di*:

(78) a. Aaming sik-zo   (***mzi**) saam-go (^{OK}**mzi**) matje

Aaming eat-PERF MZI   three-CL MZI   what

'Aaming ate some three units of something.'

b. Aaming sik-zo   (***mzi**) Aafan-ge   (^{OK}**mzi**) matje

Aaming eat-PERF MZI   Aafan-GEN MZI   what

'Aaming ate some Aafan's thing.'

The same is observed with prepositional phrases. When a PP contains *wh*-expression, *mzi* must occur PP-internally, instead of PP-externally.

(79) Aaming baai-zo   jat-bun syu   {a. ??**mzi** hai bindou/ b. hai **mzi** bindou}

Aaming put-PERF one-CL book   MZI   at   where   at   MZI where

'Aaming put a book on somewhere.'

Note that the unacceptability of these sentences does not follow from the selectional requirement of *mzi* as discussed in §2.2. This is because *three-what*, *Aafan's what* and *at where* also potentially invoke alternative sets for evaluation (i.e., the set of alternatives may *expand* until it meets a matching operator, following Rooth (1985), i.a.). Thus the strict adjacency between *mzi* and its associates is not required by semantic requirements. I therefore suggest that *mzi* might have fused with *wh*-expressions.

5.3   On the source of the ignorance component

The suggested grammaticalization path of *mzi* receives independent support from the Gifu variety of Japanese, where a verb is fused with *wh*-expressions to form epistemic indefinites. Importantly, the epistemic indefinites so formed share a common property that differentiate them from other cases,

---

(i) Aaming tai-zo   **mzi** di   matje
     Aaming read-PERF MZI CL.PL what
     'Aaming read something.'

It should however be noted that, replacing the classifier *di* with other classifiers significantly degrades the sentence and there is strong preference to the classifier-*mzi* ordering. I leave this issue to future research.

namely, the non-cancelability of the ignorance component.

Let us briefly review the epistemic indefinites in the Tokyo variety of Japanese. As argued in Alonso-Ovalle and Shimoyama ([2014]), the ignorance component of *WH-ka* in (Tokyo) Japanese is arguably a conversational implicature as it is cancelable and reinforceable, as illustrated below.

(80)   Tokyo variety of Japanese

Ken-wa   [dare-**ka**] gengogaku-no   gakusei-to   kekkonshita. jitsuwa dare-da-ka
Ken-TOP who-KA   linguistics-GEN student-with married   in.fact who-COP-Q

  shitteru.
  know

'Ken married a linguistics student. In fact, (I) know who it is.'   (p.14)

Interestingly, in Gifu variety, an additional morpheme *syan* can be attached to a WH-*ka* expression. Morphology-wise, *syan* is presumably a phonologically reduced form of *shi + ran* "know + not".[43] While the sentence in ([81]) also conveys the ignorance component of the speaker in a similar way as ([80]), it cannot be canceled.

(81)   Gifu variety of Japanese

Ken-wa   [dare **ka-syan**]-to   kekkonsi-tot-ta kedo, (#boku-wa zituwa dare-da-ka
Ken-TOP who KA-SYAN-with married   but   I-TOP   actually who-COP-KA

  sit-to-ru).
  know

'Ken married someone. In fact, I know who it is.'   (p.c. Teruyiki Mizuno)

Relevant to the current discussion is that what we observe for Cantonese *mzi* is by no means some idiosyncratic development of a certain lexical item in Cantonese; instead it represents a more general emergent pattern of epistemic markers. Furthermore, the Cantonese *mzi*-indefinites and the *WH-ka-syan* in Gifu variety reflect a close link between the morphological makeup of epistemic indefinites and the nature of the ignorance component. I stress that the particular properties of the ignorance

---

43. I thank Ken Hiraiwa and Teruyuki Mizuno for discussions.

component in *mzi*-indefinites can be attributed to the fact that *mzi* originates from an attitude verb, whose the lexical meaning becomes a non-at-issue content. This differentiates *mzi* from many other cases of epistemic indefinites whose markers do not originate as a verb and provides a partial explanation on why epistemic indefinites across languages display non-uniform properties.

6   Conclusions

In this paper, based on in-depth investigation of the string *mzi-WH* in Cantonese, I proposed that *mzi* instantiates a choice functional usage which is associated with a conventional implicature of the speakers' ignorance. I argued that *mzi* is a binder of *wh*-expressions and disjunctive expressions. It denotes a choice function and contributes to an existential interpretation of these elements. I also suggested that the ignorance component conveyed by *mzi* is best characterized as a conventional implicature in the sense of Potts (2005). It represents another possible source of an ignorance component that is less discussed in the study of epistemic indefinites. I also discussed a potential grammaticalization path of the choice functional *mzi*, which reflects a not unusual pattern across languages. It is hoped that the current study will stimulate further research on both the interpretations of *wh*-expressions in Cantonese and the study of how (speakers') ignorance is encoded in natural languages.

References

Aloni, Maria. 2001. "Quantification under Conceptual Covers." PhD diss., University of Amsterdam.

Aloni, Maria. 2007. "Free choice, modals, and imperatives." *Natural Language Semantics* 15 (1): 65–94.

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. 2006. "Disjunction in alternative semantics." PhD diss., University of Massachusetts Amherst.

Alonso-Ovalle, Luis, and Paula Menéndez-Benito. 2003. "Some Epistemic Indefinites." In *Proceedings of NELS33,* edited by Makoto Kadowaki and Shigeto Kawahara, 1–12.

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

Alonso-Ovalle, Luis, and Paula Menéndez-Benito. 2013. "Two Views on Epistemic Indefinites." *Linguistics and Language Compass* 7 (2): 105–122. doi:10.1111/lnc3.12009.

Alonso-Ovalle, Luis, and Paula Menéndez-Benito. 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): 1–56.

Cable, Seth. 2010. *The Grammar of Q: Q-Particles, Wh-Movement and Pied-Piping.* Oxford: Oxford University Press.

Charlow, Simon. 2019. "The scope of alternatives: indefiniteness and islands." *Linguistics and Philosophy.*

Chen, Sihwei, Vera Hohaus, Rebecca Laturnus, Meagan Louie, Lisa Matthewson, Hotze Rullman, Ori Simchen, Claire K. Turner, and van der Klok Jozina. 2017. "Past possibility cross-linguistically: Evidence from twelve languages." *Modality across syntactic categories:* 236–287.

Cheung, Lawrence Yam-Leung. 2015. "Uttering the unutterable with wh- placeholders." *Journal of East Asian Linguistics* 24:271–308. doi:10.1007/s10831-014-9130-x.

Chierchia, Gennaro. 2004. "A Semantics for Unaccusatives and its Syntactic Consequences." In *The Unaccusativity Puzzle: Explorations of the Syntax-Lexicon Interface,* edited by Artemis Alexiadou, Elena Anagnostopoulou, and Martin Evereart, 22–59. Oxford: Oxford University Press.

Chierchia, Gennaro. 2013. *Logic in grammar: polarity, free choice, and intervention.* Oxford: Oxford University Press.

Chierchia, Gennaro, and Hsiu-Chen Daphne Liao. 2015. "Where d Chinese *wh*-item fit?" In *Epistemic Indefinites: exploring modality beyond the verbal domain,* edited by Luis Alonso-Ovalle and Paula Menéndez-Benito, 31–59. Oxford: Oxford University Press.

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.

Dawson, Virginia. 2020. "Outscoping the directive force of imperatives." *Linguistic Inquiry:* 1–13.

Endriss, Cornelia. 2009. *Quantificaitonal topics.* Dordrecht: Springer.

Erlewine, Michael Yoshitaka. 2017. "Why the null complementizer is special in complementizer- trace effects." In *A pesky set: Papers for David Pesetsky,* edited by Claire Halpert, Hadas Kotek, and Coppe van Urk, 371–380. MIT Working Papers in Linguistics.

Farkas, Donka F. 1995. "Specicity and Scope." In *Actes du Premier Colloque Langues and Grammaire,* edited by L. Nash and G. Tsoulas, 119–137. Paris: University of Paris VIII.

Fintel, Kai von. 2004. "Would You Believe It? The King of France is Back! (Presuppositions and Truth-Value Intuitions)." In *Descriptions and beyond,* edited by Marga Reimer and Anne Bezuidenhout, 315–341. Oxford: Oxford University Press.

Fodor, Janet Dean, and Ivan Sag. 1982. "Referential and Quantificational Indefinites." *Linguistics and Philosophy* 5:355–398.

Gazdar, Gerald. 1979. *Pragmatics: implicature, presupposition, and logical form.* New York: Academic Press.

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

Geurts, Bart. 2000. "Indefinites and choice functions." *Linguistic Inquiry* 31 (4): 731–738.

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. doi:10.1016/j.lingua.2012.12.005.

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.

Hamblin, Charles. 1973. "Questions in Montague grammar." *Foundations of Language* 10 (1): 41–53.

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

Heim, Irene, and Angelika Kratzer. 1998. *Semantics in generative grammar.* Malden, MA: Blackwell.

Hopper, Paul J. 1991. "On some principles of grammaticalization." In *Approaches to grammaticalization,* edited by Elizabeth Closs Traugott and Bernd Heine, 17–36. Amsterdam: John Benjamins.

Horn, Laurence R. 1972. "On the semmantic properties of logical operators in English." PhD diss., University of California, Los Angeles. doi:10.2307/306769.

Horn, Laurence R. 2001. *A natural histotry of negation.* Standord, CA: CSLI Publications.

Huang, C.-T. James. 1987. "Existential sentences in Chinese and (in)definiteness." In *The Representation of (In)definiteness,* edited by Eric J Reuland and Alice G B ter Meulen, 226–253. Cambridge, MA: MIT Press.

Jayez, Jacques, and Lucia M. Tovena. 2006. "Epistemic Determiners." *Journal of Semantics* 23:217–250. doi:10.1093/jos/ffl002.

Johnson, Kyle. 2013. "Licensing Ellipsis." *Studies in Chinese Linguistics* 34 (2): 72–97.

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.

Kim, Ji-yung. 2003. ""Intermediate Scope" in (Mandarin) Chinese." In *Proceedings of SALT 13,* edited by R. Young and Y. Zhou, 162–179. Ithaca, NY: Cornell Linguistics Club.

Kishimoto, Hideki. 2005. "WH-in-situ and Movement in Sinhala Questions." *Natural Language and Linguistic Theory* 23 (1): 1–51.

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.

Lakoff, George. 1974. "Syntactic Amalgams." In *Papers from the 10th Regional Meeting of the Chicago Linguistic Society,* edited by Michael Galy, Robert Fox, and Anthony Bruck, 321–344.

Lee, Man Ki Theodora. 2014. "Negative *wh*-quantifiers in Cantonese: a syntactic proposal and an investigation of adult English-Cantonese interlanguage." PhD diss., University of York.

Lee, Tommy Tsz-Ming, and Hok-Yuen Wong. 2018. "Epistemic indefinites and reportative indefinites in Cantonese." In *Proceedings of the 32nd Pacific Asia Conference on Language, Information and Computation (PACLIC-32).* Hong Kong: Association for Computational Linguistics.

Lee, Tommy Tsz-Ming, and Ka-Fai Yip. 2021. "Raising, phase unlocked." In *Proceedings of the 38th West Coast Conference on Formal Linguistics,* edited by Rachel Soo, 274–282. Somerville, MA: Cascadilla Proceedings Project.

Li, Audrey Yen-Hui. 1992. "Indefinite *Wh* in Mandarin Chinese." *Journal of East Asian Linguistics* 1 (2): 12–155.

Lin, Jo-Wang. 1998. "On Existential Polarity wh-phrases in Chinese." *Journal of East Asian Linguistics* 7 (3): 219–255.

Liu, Mingming, and Yanyan Cui. 2019. "Situating Mandarin wh-indefinites in the typology of modal indefinites." *Journal of Foreign Languages* 42 (3): 26–37.

Liu, Mingming, and Yu'an Yang. 2021. "Modal wh-indefinites in Mandarin." In *Proceedings of Sinn und Bedeutung 25,* edited by P. G. Grosz, L. Marti, H. Pearson, Yasutada Sudo, and S. Zobel, 581–599. University College London / Queen Mary University of London.

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.

Ott, Dennis. 2016. "Ellipsis in appositives." *Glossa* 1 (1): 1–46. doi:10.5334/gjgl.37.

Pesetsky, David. 1987. "Wh-in-Situ: Movement and Unselective Binding." In *The Representation of (In)definiteness,* edited by Eric J Reuland and Alice G B ter Meulen, 98–129. Cambridge, Massachusetts: MIT Press.

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

Potts, Christopher. 2007. "The expressive dimension." *Theoretical Linguistics* 33 (2): 165–198.

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

Rooth, Mats. 1985. "Association with Focus." PhD diss., University of Massachusetts, Amherst.

Ruys, E G. 1992. "The scope of indefinites." PhD diss., Utrecht Univeristy.

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

Safir, Ken. 1982. "Syntactic Chains and the Definiteness Effect." PhD diss., Massachusetts Institute of Technology.

Shanon, Benny. 1976. "Semantic and syntactic features in the context of dreams." *Linguistic Inquiry* 7 (3): 518–519.

Shimoyama, Junko. 2006. "Indeterminate Phrase Quantification in Japanese." *Natural Language Semantics* 139–173 (14): 2.

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.

Simons, Mandy. 2005. "Semantics and Pragmatics in the Interpretation of 'or'." In *Proceedings of SALT XV,* edited by Effi Georgala and Jonathan Howell, 205–222. Cornell University, Ithaca, NY: CLC Publications.

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. doi:10.1093/acprof:oso/9780199665297.001.0001.

Strawson, P.F. 1974. *Subject and Predicate in Logic and Grammar.* London: Routledge.

Sudo, Yasutada. 2010. "Wh-ka Indefinites in Japanese," 1–19. Handout distributed at Workshop on Epistemic Indefinites, University of Göttingen on June 11, 2010.

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

Wong, Hok-Yuen. 2018. *Wh-doublets in Cantonese.* MA thesis, The Chinese University of Hong Kong.

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.