

Epistemic Indefinites and Reportative Indefinites in Cantonese

Tommy Tsz-Ming Lee

University of Southern California
 tszmingl@usc.edu

Hok-Yuen Wong

The Chinese University of Hong Kong
 wonghyo@cuhk.edu.hk

Abstract

This paper examines the syntactic and semantic properties of the two understudied uses of *wh*-expressions in Cantonese, in the form of *m-zi-wh* ‘not-know-wh’ and *wh-wh* (wh-doublets). We argue that they are, respectively, *Epistemic Indefinites* (EIs) and *Reportative Indefinites* (RIs). We suggest that the speaker’s ignorance associated with EIs and reportative implication associated with RIs, are conventional implicatures. Building on their distributional similarities, we propose a unified syntactic and semantic analysis for both EIs and RIs in Cantonese.

1. Introduction

The interpretation of *wh*-expressions in Cantonese varies with environments, largely patterning with Mandarin (Chierchia and Liao, 2015, i.a.).

(1) Interrogative

Aaming sik-zo **matje**
 Aaming eat-PERF **what**
 ‘What did Aaming eat?’

(2) Negative (Lee 2014)

Aaming mou sik **matje**
 Aaming not eat **what**
 ‘Aaming didn’t eat anything.’ /
 ‘Aaming ate only a few things.’

(3) Universal

Aaming **matje** dou sik
 Aaming **what** all eat
 ‘Aaming eats everything.’

A complication lies in the existential use of Cantonese *wh*-expressions, which cannot be used in exactly the same way as Mandarin ones.

(4) Existential

a. *Aaming houci sik-zo **matje** [C]
 Aaming seem eat-PERF **what**

Int.: ‘Aaming seems to have eaten something.’

b. Aaming haoxiang chi-le **shenme** [M]
 Aaming seem eat-PERF **what**

Instead of bare *wh*-expressions, Cantonese adopts a numeral expression (i.e. (one)-CL NP) to express (4a).

(5) Aaming houci sik-zo (**jat**) **joeng je**
 Aaming seem eat-PERF **one kind thing**

It is, however, premature to conclude that *wh*-expressions in Cantonese cannot be used existentially in any case. Truth-conditionally, both (6) and (7) entail (5).

(6) *Wh*-expressions in *m-zi-wh* sequence

Aaming houci sik-zo **m-zi matje**
 Aaming seem eat-PERF **M-ZI what**
 ‘Aaming seems to have eaten something_u.’¹

(7) *Wh*-expressions in the form of *wh*-doublets

Aaming houci sik-zo **matje-matje**
 Aaming seem eat-PERF **what-what**
 ‘Aaming seems to have eaten something_m.’

However, these uses convey something more than a canonical indefinite expression. For example, (6) conveys the speaker’s ignorance on what Aaming seems to have eaten, while (7) suggests that the thing that Aaming seems to have eaten is previously mentioned to the speaker.

The goal of this paper is to examine the syntax and semantics of these uses of *wh*-expressions in (6) and (7). We first argue that the *wh*-expression in (6) is used as an *Epistemic Indefinite* (hereafter EI; in the sense of Alonso-Ovalle and Menéndez-Benito (2015)), whereas the *wh*-expression in (7) as a *Reportative Indefinite* (hereafter RI; cf. Quotational Indefinite discussed in Koev (2017)), respectively (§2.1 & §2.2). We further argue that the “additional meaning” associated with EIs and RIs are conventional implicatures (Potts 2005) (§2.3). In §3, we reveal three various distributional similarities of EIs and RIs. We give

¹ The subscripts u and m in the translation lines stand for ‘unknown to the speaker’ and ‘mentioned to the speaker previously’ respectively.

our syntactic and semantic proposal for EIs and RIs in §4, followed by discussion on two miscellaneous issues in §5. We conclude the paper in §6.

2. Indefinites with additional implications

2.1. *M-zi-wh* as Epistemic Indefinites

We argue that the *m-zi-wh* sequence (as in (6)) is used as an EI.² By EIs, we mean indefinites that “signal ignorance on the part of the speaker, thereby conveying information about her epistemic state” (Alonso-Ovalle and Menéndez-Benito, 2015:2). In other words, EIs are used with speaker’s failure in identification of the referent. We call it the ignorance component in the rest of the paper.

The ignorance component can be detected by the infelicity of continuations that encodes speakers’ knowledge on the matter at issue, such as the “namely” test in (8b) and the “guess-what” test in (8c). Both of them fails to be a felicitous continuation of (8a).

- (8) a. Aaming maai-zo **m-zi bin-bun** syu
Aaming buy-PERF **M-ZI which-CL** book
‘Aaming bought some book_m.’
b. # ...zikhai Hunglaumung
namely Dream.of.the.red.chamber
‘...namely, *Dream of the Red Chamber*.’
c. # ...nei gu-haa hai bin-bun
you guess-try be which-CL
‘...guess which one it is.’

On the contrary, in (9), where the EI is replaced by a numeral expression, both continuations of (8b) and (8c) are felicitous.

- (9) Aaming maai-zo jat-bun syu
Aaming buy-PERF one-CL book
‘Aaming bought a book.’

2.2. *Wh*-doublets as Reportative Indefinites

Next, we suggest that *wh*-doublets are used as RIs in Cantonese.³ By RIs, we mean indefinites that trigger reportative implication. Functionally, they are used to substitute some content that is previously mentioned to the speaker (cf. Koev, 2017). They can be licensed through grammatical means, e.g. under a verb of saying (10) or a

hearsay evidential particle *wo5* (11), or a proper reportative discourse (12).

- (10) Verbs of saying: e.g. *waa* ‘say’
??(Aaming waa) keoi maai-zo
Aaming say he buy-PERF
bin-bun-bin-bun syu
which-CL-which-CL book
‘Aaming said he bought some book_m.’
(11) Hearsay evidential particle *wo5*
Aaming maai-zo **bin-bun-bin-bun**
Aaming buy-PERF **which-CL-which-CL**
syu ??(*wo5*)
book SP_{hearsay}
‘Aaming wants to buy some book_m.’
(12) Reportative discourse
A: Aaming waa keoi ceot-zo-gaai
Aaming say he hang.around-PRT
‘Aaming said he hung around
somewhere.’
B: Gam keoi zong gong-zo matje
then he also say-PRT what
‘Then what did he also say?’
A: Keoi soeng maai **bin-bun-bin-bun**
he want buy **which-CL-which-CL**
syu lol
book SP_{non-hearsay}
‘He wants to buy some book_m.’

Unsurprisingly, an out-of-the-blue context, for example, the beginning of a story in (13), renders the use of RIs infelicitous.

- (13) [The speaker begins to tell a story.]
Cungcin, Aaming maai-zo
once Aaming buy-PERF
bin-bun-bin-bun syu
which-CL-which-CL book
‘Once, Aaming bought some book_m.’

The reportative implication is removed once the RI is replaced by a numeral expression. Compare (10) and (14). In (10), the speaker must have heard about the book titles, for example, from Aaming in some previous conversation; while it need not be the case in (14).

- (14) Aaming waa keoi maai-zo jat-bun syu
Aaming say he buy-PERF one-CL book
‘Aaming said he bought a book.’

² EIs in Cantonese are a sequence of negation *m* ‘not’, a predicate *zi* ‘know’ and a *wh*-expression, whereas EIs in European languages usually bear a determiner- or quantifier-like expression, such as the Spanish *algún* and the German *irgendein*.

³ We coin the term RIs along the line of Koev’s (2017) Quotational Indefinites (QIs), but use it as a cover term for QIs, since, as we put it, they range over not only *expressions*, but also *entities*.

2.3. Conventional implicatures in EIs and RIs

In the literature pertinent to EIs, the ignorance component of EIs can be attributed to *conversational implicature* based on Grice's (1975) reasoning (i.e. Maxim of Quantity), as suggested in Kratzer and Shimoyama (2002) for German, Alonso-Ovalle and Menéndez-Benito (2010) for Spanish, and Alonso-Ovalle and Shimoyama (2014) for Japanese. However, Cantonese data below shows that the ignorance component in EIs cannot be cancelled or be reinforced without redundancy, as suggested by the infelicity of (15b) and (15c) as continuations of (15a).

- (15) a. Aaming gin-dou **m-zi bingo**
 Aaming see-ACH **M-ZI who**
 'Aamingsaw some person_u.'
 b. # ...keisat ngo zi hai bingo
 actually I know be who
 '...Actually, I know who he is.'
 c. # ...daanhai ngo m-zi hai bingo
 but I not-know be who
 '...But I don't know who he is.'

Also, the ignorance component does *not* disappear under downward entailing environments. (16) asserts the existence of an assignment that Aaming didn't submit, one that the speaker cannot identify (i.e. the ignorance component remains). (16) does not mean that Aaming didn't submit *any* assignment (i.e. the ignorance component disappears), an inference observed in German and Spanish EIs (e.g. Alonso-Ovalle and Menéndez-Benito, 2010).

- (16) Aaming mou-gaau **m-zi bin-fan**
 Aaming not-submit **M-ZI which-CL**
 gungfo
 assignment
 a. 'Aaming didn't submit some assignment_u.'
 b. # 'Aaming didn't submit any assignment.'

Following the same line of reasoning, the reportative implication associated with RIs is also unlikely to be a conversational implicature. It is unclear how it may be related to Gricean reasoning to start with. In addition, the continuation in (17), which denies the previous mentioning of the book title (i.e. an attempt to cancel the reportative implication), is infelicitous to (10) and (11) above, suggesting the uncancellability of the reportative implication.

- (17) ...daan keoi keisat matje syumeng
 but he actually what book.title

dou mou-gong-gwo
 all not-say-EXP
 '...But he didn't mention any book title.'

Similarly, the reportative implication remains even if the RI scopes under negation, as in (18). We conclude that both the ignorance component and the reportative implication are not an instance of conversational implicature.

- (18) Keoi mou maai **bin-bun-bin-bun**
 he not.have buy **which-CL-which-CL**
 syu wo5
 book SP_{hearsay}
 a. 'He didn't buy a book_m.'
 b. # 'He didn't buy any book.'

We propose, instead, that both the additional meanings associated with EIs and RIs are indeed *conventional implicatures* (hereafter CIs) in Potts' (2005, 2015) sense, defined as follows:

- (19) Meaning *p* is a *CI* of phrase *S* iff
 a. *p* is a conventional (encoded) property of a lexical item or construction in *S*;
 b. *p* is entailed by *S*; and,
 c. *p*'s truth or falsity has no effect on the *at-issue* content of *S*. (Potts 2015:(17))

To see how EIs and RIs fit into the picture depicted in (19a) and (19b), the ignorance component in EIs is encoded by *m-zi*. In other words, the ignorance component is *entailed* in EIs. There is no way to use a Cantonese EI without expressing speakers' ignorance. On the other hand, the reportative implication in RIs is encoded by *wh*-doublets. We agree with Xu (2010) in that *wh*-doublets are related to reportativity, implying the existence of a reporting event. It explains why RIs are only licensed in reportative contexts (§2.2). Reportative implication is also *entailed* in RIs.

With regard to (19c), the logical independence between *at-issue* content and CI can be verified by "yes, but..." test in (20)-(21) (Karttunen and Peters, 1979; Potts, 2005). The response "yes, but..." disputes only the CI content, but not the *at-issue content*. In particular, the "yes" in B's utterance agrees with the truth value of the *at-issue* content (i.e. Aaming ate something), while the correction made by "but"-clause focuses on CI only.

- (20) A: Aaming sik-zo **m-zi matje**
 Aaming eat-PERF **M-ZI what**
 'Aaming has eaten something_u.'
 B: Hai aa3, batgwo nei zi keoi
 yes SP but you know he

sik-zo matje gaak3
eat-PERF what SP
'Yes, but you know what he has eaten.'

- (21) A: Aaming sik-zo **matje-matje** wo5
Aaming eat-PERF **what-what** SP
'Aaming has eaten something_m.'
B: Hai aa3, batgwo mou-jan gong-gwo
yes SP but no-person say-EXP
hai matje jesik bo3
be what food SP
'Yes, but no one mentioned the food he has eaten.'

We conclude that both the ignorance component and the reportative implication are highly conventionalized and are best captured under a CI analysis (cf. a similar approach to EIs in French, i.e. *quelque* (Jayez and Tovená, 2008, 2013), and RIs in Bulgarian (Koev, 2017)).⁴

3. Distributional similarities between EIs and RIs

3.1. External syntax

As mentioned in §1, both EIs and RIs are indefinites. Their distribution pattern with one-CL-NP indefinites in *jau*-construction. The *jau*-construction shows Definiteness Effect (Huang, 1987). That is, only indefinite NPs, but not definite NPs, can appear after an existential marker *jau* in *jau*-construction.

- (22) Jau {**m-zi bin-fan** gungfo /
have **M-ZI which-CL** assignment
bin-fan-bin-fan gungfo /
which-CL-which-CL assignment
jat-fan gungfo},
one-CL assignment
Aaming m-geidak zou wo5
Aaming not-remember do SP
'There is some assignment that Aaming forgot to do.'

Huang also suggests that the construction with an experiential suffix attached to the main verb, like (23) below, always has an indefinite after the suffix. Both EI and RI can appear in the same position as one-CL-NP.

- (23) Ngo gaau-gwo {**m-zi bingo** hoksaang /
I teach-EXP **M-ZI who** student
bingo-bingo hoksaang / jat-go hoksaang}
who-who student one-cl student

hou congming wo5
very clever SP
'I have the experience of teaching a student who is very clever.'

3.2. Beyond the nominal domain

M-zi can combine with *wh*-adverbials and *wh*-doublets are also possible with *wh*-adverbials. Similar ignorance component and reportative implication are found in these sequences.

- (24) Aaming **m-zi dimgaai** cidou
Aaming **M-ZI why** come.late
'Aaming came late for some reason_u.'
(25) Aaming **dimgaai-dimgaai** cidou wo5
Aaming **why-why** come.late SP
'Aaming came late for some reason_m.'

3.3. Strict adjacency

In EIs, *m-zi* and the following *wh*-expression are always adjacent.

- (26) a. * Aaming sik-zo **m-zi**
Aaming eat-PERF **M-ZI**
[Num-CL saam-go] **matje**
three-CL **what**
Int.: 'Aaming ate some three things_u.'
b. * Aaming sik-zo **m-zi**
Aaming eat-PERF **M-ZI**
[Possessor Aafan ge] **matje**
Aafan POSS **what**
Int.: 'Aaming ate some Aafan's thing_u.'

A complication here is that there are cases where *m-zi* can occur sentence-initially (27a) or preverbally (27b), giving almost identical meaning. However, *m-zi* here is arguably different from its use in EI. See §5.1 for discussion.

- (27) a. **M-zi** Aaming sik-zo **matje** (le1)
M-ZI Aaming eat-PERF **what** SP
b. Aaming **m-zi** sik-zo **matje** (le1)
Aaming **M-ZI** eat-PERF **what** SP

On the other hand, the adjacency requirement for RIs is obvious. The two *wh*-expressions cannot be intervened by any other elements, e.g. preposition *hai* 'at' (28).

- (28) Aaming hai **bindou** (*hai) **bindou**
Aaming at **where** at **where**
maai-zo syu wo5
buy-PERF book SP
'Aaming bought books at somewhere_m.'

⁴ We do not further distinguish between CI and presupposition, for its irrelevance here. But see discussion in Potts (2005) and Koev (2017).

Given that RIs in some languages preferably denote a plural reading (e.g. Mandarin (Yu, 1964; Bai, 2008)), one may treat the Cantonese RI as a coordination of two *wh*-expressions with a silent *tung* ‘and’. However, this plurality requirement is found to be void in Cantonese, as singular reading is available (Wong, 2018). Inserting an overt coordinator in between the RI forces a plural reading:

- (29) Aaming hai **bindou** tung **bindou**
 Aaming at **where** and **where**
 maai-so syu wo5
 buy-PERF book SP
 ‘Aaming bought books at two different places.’

An account adopting coordinate structure is therefore implausible in Cantonese RIs. We conclude that the two *wh*-expressions in RIs must be adjacent to each other.

4. Analysis

The parallels reported in §2 and §3 motivate a unified syntactic account for EIs and RIs. We propose that their use as indefinites is the *at-issue* content of the sentence (§4.1), while the conventional implicatures associated with EIs and RIs are best captured under a CI analysis in the sense of Potts (2005) (§4.2).

4.1. The *at-issue* content in EIs and RIs

The proposal here assumes the framework of focus semantics (Rooth, 1985, 1992). The basic idea is on the same line as Hagstrom (1998), Beck (2006) and Cable (2010). For the interpretation of *wh*-expressions, we assume that *wh*-expressions are semantically deficient, and come with focus semantic value but not normal semantic value (i.e. unbound variables). For example, the semantics of *matje* ‘what’ is two-layered:

- (30) a. Normal semantics
 $\llbracket \text{matje} \rrbracket = \text{undefined}$
 b. Focus semantics
 $\llbracket \text{matje}_F \rrbracket^F = \{x: x \notin \text{human}\}$

We propose that the internal structure of both EIs and RIs are the same, schematized as (31), where *m-zi* in EI and the first *wh*-expression in RI, abbreviated as *WH*, head a functional position Y, taking a *wh*-expression as their complement. *M-zi* and *WH* project a YP. Note that the category of Y depends on the category of their complement which can be a *wh*-nominal or a *wh*-adverbial.

- (31) The syntactic structure of EIs and RIs
 $\llbracket \text{YP } m\text{-}zi_i / WH_i \llbracket \text{XP } wh\text{-expression} \rrbracket \rrbracket$

For the semantics of *m-zi* and *WH*, we propose that they are *focus-sensitive operators* in the sense of Rooth (1985, 1992), which take as input the focus semantics of its complement (Beck, 2006). Semantically, they are *variables over choice functions* (Cable, 2010, i.a.). They take a (focus) set as its argument and return a member of the set as its value. They are, and indeed must be, bound by a higher operator (i.e. the existential closure of choice functions). Their semantics are given as (32). Following Beck (2006), we assume a syncategorematic rule for these focus-sensitive operators as in (33).

- (32) Semantics of *m-zi* and *WH* (first version)
 $\llbracket m\text{-}zi_i / WH_i \rrbracket^g = g(i) \in D_{cf}$

- (33) A special compositional rule for EIs and RIs
 $\llbracket m\text{-}zi_i / WH_i \text{ XP} \rrbracket = \llbracket m\text{-}zi_i / WH_i \rrbracket (\llbracket \text{XP} \rrbracket^F)$

For RIs, we suggest that the phonetic features of *WH* is acquired during derivation. In particular, it copies whatever phonetic features of its complement at the PF component (resulting in reduplication, see §5.2 for complications in formation of RIs).

As an illustration, the structure of (6) and (7) can be represented by (34), setting aside the semantics of the adverb *houci* ‘seem’ and the particle *wo5*. We assume that the existential closure of choice function variables takes widest scope.

- (34) $\llbracket \text{TP}_2 \exists f[\text{TP}_1 \text{ Aaming} \llbracket \text{VP } sik\text{-}zo \llbracket \text{DP}_2 m\text{-}zi_i / WH_i \llbracket \text{DP}_1 \text{ matje} \rrbracket \rrbracket \rrbracket \rrbracket$

Let us first compute the semantics for the EI and RI in (34):

- (35) $\llbracket \text{DP}_2 \rrbracket = (\text{by (33)})$
 $\llbracket m\text{-}zi_i / WH_i \rrbracket (\llbracket \text{matje}_F \rrbracket^F) = (\text{by (30), (32)})$
 $g(1) \in D_{cf}(\{x: x \notin \text{human}\}) =$
 $f(\{\text{the apple, the cake, the cookies, ...}\})$

Assume the semantics of *sik-zo* ‘ate’ to be (36):

- (36) $\llbracket sik\text{-}zo \rrbracket = \lambda y \lambda x. x \text{ has eaten } y$

The semantics for (34) can be computed as follows:

- (37) $\llbracket \text{TP}_2 \rrbracket = \exists f[\llbracket \text{VP} \rrbracket(\text{Aaming})] =$
 $\exists f[\llbracket sik\text{-}zo \rrbracket(\llbracket \text{DP}_2 \rrbracket)(\text{Aaming})] = (\text{by (35)})$
 $\exists f[\llbracket sik\text{-}zo \rrbracket(f(\{\text{the apple, the cake, the cookies, ...}\}))(\text{Aaming})] = (\text{by (36)})$
 $\exists f[\text{Aaming has eaten } f(\{\text{the apple, the cake, the cookies, ...}\})]$

(Read as: ‘There is a way of choosing an element from the set such that Aaming has eaten the element so chosen.’)

4.2. The CIs in EIs and RIs

Recall that the ignorance component in EIs and the reportative implication in RIs are conventional implicatures, rather than conversational implicatures (§2.3). Along the view of Karttunen and Peters (1979) and Potts (2005), these CI properties motivate a multidimensional theory of meaning, in which the semantics of *at-issue* content and that of CIs are derived separately. Since a sentence may have multiple expressions associated with different CIs, we can imagine an n -ary tuple of truth values, for any integer n . For simplicity, we assume that only one CI expression is available in a sentence. Following Koev (2017), the meaning of a sentence is usually two-dimensional, formalized as an ordered pair with the *at-issue* content constituting the first component and the CI the second:

$$(38) \llbracket \cdot \rrbracket = \langle \{at\text{-issue content}\}, \{CI\} \rangle$$

A full formal implementation of two-dimensional semantics can be found in Potts (2005) and Koev (2017). Here we only give the semi-formal semantics for EIs and RIs. Based on (32) and the multidimensional analysis of CI, we propose the following semantics for EIs (39) and RIs (40), which maintain the parallelism as existential indefinites and capture their difference with regard to CIs:

(39) Semantics for *m-zi* (final version)

$$\llbracket m\text{-}zi \rrbracket^g = \langle g(i) \in D_{cf}, \text{fail.to.identify}(the\ speaker, g(i)) \rangle$$

(40) Semantics for *WH* (final version)

$$\llbracket WH \rrbracket^g = \langle g(i) \in D_{cf}, \exists s. \text{mention.to.the.speaker}(s, g(i)) \rangle$$

(where s is some other speaker)

For concreteness, the full semantics of (6) and (7) can be given as, respectively:

$$(41) \llbracket (6) \rrbracket^g = \langle \exists f[\text{Aaming seems to have eaten } f(\{\text{the apple, the cake, the cookies, ...}\})], \text{fail.to.identify}(the\ speaker, f) \rangle$$

$$(42) \llbracket (7) \rrbracket^g = \langle \exists f[\text{Aaming seems to have eaten } f(\{\text{the apple, the cake, the cookies, ...}\})], \exists s. \text{mention.to.the.speaker}(s, f) \rangle$$

5. Miscellaneous issues

This section discusses two puzzles stemming from different grammatical status of preverbal and postverbal *m-zi* (§5.1) and different variant forms of RIs (§5.2).

5.1. Preverbal and postverbal *m-zi*

M-zi can appear in different positions in a sentence: sentence-initial position as in (43a), post-subject position as in (43b) and the post-verbal position as in (43c).

- (43) a. **M-zi** Aaming sik-zo **matje**
M-ZI Aaming eat-PERF **what**
 b. Aaming **m-zi** sik-zo **matje**
 Aaming **M-ZI** eat-PERF **what**
 c. Aaming sik-zo **m-zi** **matje**
 Aaming eat-PERF **M-ZI** **what**
 ≈ ‘Aaming ate something_u.’

Although the meaning difference is subtle, we argue that *m-zi* in (43a) and (43b) (hereafter preverbal use of *m-zi*) are (matrix) predicates (cf. Yap & Chor 2014), whereas that in (43c) (hereafter post-verbal use of *m-zi*) is not. According to our analysis in §4, the post-verbal *m-zi* is a focus-sensitive operator. We present three arguments in favor of this claim.

The first argument concerns subject recoverability. Yap and Chor (2014) suggests a *pro*-drop analysis of (43a), where the subject ‘I’ is null and they derive (43b) from (43a) by topicalizing the embedded subject (i.e. ‘Aaming’). Since *m-zi* is a predicate in the pre-verbal usage, the dropped subject ‘I’ (i.e. the external argument of *m-zi*) can be phonetically recovered:

- (44) a. (Ngo) **m-zi** Aaming sik-zo **matje**
 I **M-ZI** Aaming eat-PERF **what**
 b. Aaming_i ngo **m-zi** $\{?t_i / \text{OK}keoi_i\}$
 Aaming I **M-ZI** he
 sik-zo **matje**
 eat-PERF **what**

The marginality of (44b) comes from the fact that *m-zi* is grammaticalized as a (negative) attitudinal marker (via insubordination). The previous external argument becomes first person subject by default and so recovering the subject leads to redundancy and hence marginality. Resumptive pronoun may rescue (44b). Note that it is substantially better than (45), where post-verbal *m-zi* cannot take any external argument.

- (45)* Aaming sik-zo ngo **m-zi** **matje**
 Aaming eat-PERF I **M-ZI** **what**

Although they do not discuss the case of post-verbal *m-zì*, it is possible, following their line of reasoning, that grammaticalization takes place further, depriving *m-zì* of its predicate status. The external argument is entirely absorbed (and hence speaker-directed ignorance). No subject is allowed. No rescue (e.g. resumptive pronoun) is possible.

The second argument concerns the compatibility with sentence-final particle *lel*. Witness the contrast in (46):

- (46) a. **M-zì** Aaming sik-zo **matje** lel
 M-ZI Aaming eat-PERF **what** SP
 b. Aaming **m-zì** sik-zo **matje** lel
 Aaming **M-ZI** eat-PERF **what** SP
 c. * Aaming sik-zo **m-zì** **matje** lel
 Aaming eat-PERF **M-ZI** **what** SP

lel is compatible with both matrix questions and embedded questions, but not with declarative sentences.

- (47) a. Matrix question (M&Y 2011:427)
 Bindou hai lel?
 where be SO
 ‘How come?’
 b. Embedded question (M&Y 2011:410)
 Nei waa [keoi hai-m-hai zeonbou-zo
 you say he be-not-be improve-PERF
 lel]?
 SP
 ‘Do you think he has improved?’
 c. Declarative sentence
 * Keoi hai zeonbou-zo lel.
 he be improve-PERF SP
 Int.: ‘He has improved.’

It is known that *m-zì*, as a predicate, can take interrogative complement, forming embedded question and their compatibility with *lel* in (46a) and (46b) follows. The unacceptability in (46c) can be explained by suggesting that no embedded question is formed. If (46c) does not form an embedded question, *m-zì* in (46c) is not taking an interrogative complement, different from *m-zì* in (47a) and (47b). A natural explanation is that *m-zì* in (47c) is not a predicate (but a focus operator which does not form embedded question).

The last argument concerns possible continuations of the sentences in (43). (48) is a felicitous continuation of (47a) and (47b), but not (47c).

- (48) ...daan honang me-dou mou-sik
 but possibly what-all not.have-eat
 ‘...but, maybe, (he) hasn’t eaten anything.’

Although the embedded questions in (47a) and (47b) presuppose Aaming ate something, it can be *retreated* by (48). However, the infelicity of (48) as a continuation of (47c) suggests that the presupposition cannot be retreated. Intuitively, the speaker is making a contradictory statement with both (47c) and (48). It is suggestive that the proposition that Aaming ate something is not *presupposed*, but *asserted* in (47c). *M-zì* in (47a) and (47b), like other predicate *soeng-zidou* ‘wonder’ which also takes interrogative complement, does not assert the proposition expressed by its complement (i.e. embedded question does not make any assertion). The assertion made in (47c) suggests *m-zì* does not serve as a matrix predicate.

5.2. Derivation of variant forms of RIs

In addition to the proposed PF copying operation, the formation of RIs is constrained by some other conditions. Consider the following paradigm:

- (49) a. * Keoi gin-dou **bin-bin-go** wo5
 b. Keoi gin-dou **bingo-bingo** wo5
 c. Keoi gin-dou **bingo-bingo-jan** wo5
 d. ? Keoi gin-dou **bingo-jan-bingo-jan** wo5
 ‘Aaming said he saw someone_m.’

We suggested that the phonetic features of *WH* are acquired during derivation. In particular, they copy the features of its complement. Implicit in the proposal is the idea that the PF copying operation works at the minimal phrasal level only. The ungrammaticality of (49a) verifies the phrasal requirement in RI reduplication (on the assumption that *bin* ‘which’ is a D head). This proposal also correctly predicts (49b), as the whole phrase *bingo*, presumably a DP, is copied. However, (49d) appears to be a counterexample. We postulate NP ellipsis as an additional PF rule in regulating the formation of RIs: when the reduplicated phrasal form is too “heavy”, the first head noun tends to be elided. We thus have the acceptable (49c) (where NP ellipsis occurs) and the marginal (49d) (where NP ellipsis is absent). The same applies to *matje* ‘what’:

- (50) a. Keoi maai-zo bun **matje-matje** wo5
 (phrasal copying)
 b. Keoi maai-zo bun **matje-matje-syu** wo5
 (phrasal copying & NP ellipsis)
 c. ? Keoi maai-zo bun **matje-syu-matje-syu**
 wo5 (phrasal copying)
 ‘He bought some book_m.’

As for Reportative Adverbials (RAs), if we assume *wh*-adverbials are phrasal, then the pattern in (51) follows. PF Copying at non-minimal phrasal level is disallowed.

- (51) a. Keoi **dimgaai-dimgaai** maai bun syu
wo5 (phrasal copying)
b. * Keoi **dimgaai-maai-dimgaai-maai** bun
syu wo5 (non-phrasal copying)
c. * Keoi **dimgaai-maai-bun-syu-dimgaai-**
maai-bun-syu wo5
(non-minimal phrasal copying)
‘He buys a book for some reason_m.’

Another type of RI variants is *wh*-triplets (52). Mono-syllabic phrasal *wh*-expressions (i.e. *mat* and *dim*, both of which are presumably a reduced form of *matje* and *dimjoeng*) forms *wh*-triplets instead of *wh*-doublets. We regard the choice between *wh*-doublets and *wh*-triplets as a phonological one.

- (52) a. Aaming waa keoi maai-zo bun
mat-mat-(**mat**) wo5
b. Aaming waa keoi daa-dou go deoisau
dim-dim-(**dim**) wo5

In sum, the derivation process of an RI at PF is as follows:

- (53) (i) Optional syllable reduction of *wh*-expressions (only applicable to *matje* and *dimjoeng*);
(ii) PF copying operation targeting the complement of *WH*
- reduplication for disyllabic *wh*-expressions
- triplication for monosyllabic *wh*-expressions
(iii) NP ellipsis of the first head noun

6. Concluding remarks

This paper examined the syntactic and semantic properties of the two understudied uses of *wh*-expressions in Cantonese, in the form of *m-zi-wh* ‘not-know-*wh*’ and *wh-wh* (*wh*-doublets). We argued that they are, respectively, *Epistemic Indefinites* (EIs) and *Reportative Indefinites* (RIs) (§2.1 & §2.2). We suggested that the speaker’s ignorance associated with EIs and reportative implication associated with RIs, are indeed conventional implicatures (§2.3). Building on their distributional similarities (§3), we proposed a unified syntactic and semantic analysis for both EIs and RIs in Cantonese (§4), and discussed two miscellaneous issues related to EIs and RIs (§5).

Some issues remain unresolved. First, there is indeed a distributional asymmetry in possible arguments for EIs and RIs. The operator *m-zi* can take a disjunction set as its complement, but *WH* cannot. A possibility is that *m-zi* in (54a), despite its post-verbal position, is indeed a matrix predicate, but not the *m-zi* operator in our proposal, since it is compatible with SP *le1*.

- (54) a. Aaming jam-zo **m-zi seoi-ding-caa**
Aaming drink-PERF **m-zi water-or-tea**
‘Aaming drank either water or tea_u.’
b. * Aaming jam-zo
Aaming drink-PERF
seoi-ding-caa-seoi-ding-caa wo5
water-or-tea-water-or-tea SP
Int.: ‘Aaming drank either water or tea_m.’

Second, the adjacency between *m-zi* and *WH* is surprisingly strict. If we follow Hamblin (1973) where the focus set denoted by the *wh*-expressions can expand until they meet a suitable binder, we expect to see *m-zi* and *WH* occur at a distance from its *wh*-complement, but this is not the case, a property that is substantially different from the Q-particles (Cable 2010). A possibility is that *m-zi* and *WH* are affix-like elements that require phonological support from its complement. We leave these issues to future research.

Acknowledgments

We are grateful to two anonymous reviewers for comments and suggestions. We have also benefited from discussion with, according to the alphabetical order of surnames, Lawrence Cheung, Audrey Li, Rumi Pancheva, and also the FoCaL reading group. All remaining errors are ours.

References

- Alonso-Ovalle, Luis, and Paula Menéndez-Benito. 2010. Modal indefinites. *Natural Language Semantics* 18(1): 1–31.
Alonso-Ovalle, Luis, and Paula Menéndez-Benito. 2015. Epistemic indefinites: An overview. In *Epistemic indefinites: Exploring modality beyond the verbal domain*, eds. Luis Alonso-Ovalle, and Paula Menéndez-Benito, 1–30. 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*, ed.

- Robert E. Santana-LaBarge, 11–20. Somerville, MA: Cascadilla Proceedings Project.
- Bai, Yan. 2008. Xiandai hanyu daici chongdie xianxiang zhi kaocha [An investigation of the reduplication of pronouns in contemporary Chinese]. MA thesis, Central China Normal University.
- Beck, Sigrid. 2006. Intervention effects follow from focus interpretation. *Natural Language Semantics* 14: 1–56.
- Cable, Seth. 2010. *The grammar of Q: Q-particles, wh-movement, and pied-piping*. Oxford: Oxford University Press.
- Chierchia, Gennaro, and Hsiu-Chen Liao. 2015. Where do Chinese *wh*-items fit? In *Epistemic indefinites: Exploring modality beyond the verbal domain*, eds. Luis Alonso-Ovalle, and Paula Menéndez-Benito, 31–59. Oxford University Press.
- Grice, H. Paul. 1975. Logic and conversation. In *Syntax and semantics, iii*, eds. P. Cole, and J. Morgan, 43–58. New York: Academic Press.
- Hagstrom, Paul. 1998. Decomposing questions. PhD dissertation. MIT, Cambridge, MA.
- Hamblin, Charles L. 1973. Questions in Montague English. *Foundations of Language* 10: 41–53.
- Huang, James Cheng-Teh. 1987. Existential sentences in Chinese and (in)definiteness. In *The representation of (in)definiteness*, eds. Eric J. Reuland, and Alice G. B. ter Meulen, 226–253. Cambridge, MA: MIT Press.
- Jayez, Jacques, and Lucia M. Tovena. 2013. Scenarios of equivalence: The case of *quelque*. In *Different kinds of specificity across languages*, eds. Cornelia Ebert, and Stefan Hinterwimmer, 177–207. Springer.
- Karttunen, Lauri, and Stanley Peters. 1979. Conventional implicature. In *Syntax and semantics, xi: Presupposition*, eds. Choon-Kyu Oh, and David Dinneen, 1–56. New York: Academic Press.
- Koev, Todor. 2017. Quotational indefinites. *Natural Language and Linguistic Theory* 35: 367–396.
- Kratzer, Angelika, and Junko Shimoyama. 2002. Indeterminate pronouns: The view from Japanese. In the *Proceedings of the Third Tokyo Conference on Psycholinguistics (TCP)*, ed. Yukio Otsu, 1–25. Tokyo: Hituzi Syobo.
- Lee, Man-Ki Theodora. 2014. Negative *wh*-quantifiers in Cantonese: A syntactic proposal and an investigation of adult English-Cantonese interlanguage. Unpublished PhD dissertation.
- Matthews, Stephen, and Virginia Yip. 2011. *Cantonese: A comprehensive grammar* (2nd ed). London; New York: Routledge.
- Potts, Christopher. 2005. *The logic of conventional implicatures*. Oxford: Oxford University Press.
- Potts, Christopher. 2015. Presupposition and implicature. In *The handbook of contemporary semantic theory* (2nd ed.), eds. Shalom Lappin, and Chris Fox, 168–202. John Wiley & Sons.
- Rooth, Mats. 1985. Association with focus. PhD dissertation, University of Massachusetts at Amherst.
- Rooth, Mats. 1992. A theory of focus interpretation. *Natural Language Semantics* 1: 75–116.
- Wong, Hok-Yuen. 2018. *Wh*-doublets in Cantonese. Unpublished MPhil thesis, The Chinese University of Hong Kong.
- Xu, Muofan. 2010. Lun yiwendaici zhidai yongfa de chongdie [Reduplication of interrogative pronouns with referential use]. *Yuyan jiaoxue yu yanjiu [Language Teaching and Research]*: 56–62.
- 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*, eds. Elisabeth Leiss, and Werner Abraham, 219–260. Amsterdam, Netherlands: John Benjamins.
- Yu, Xiliang. 1964. Yiwendaici de chongdie yongfa [The reduplicated use of interrogative pronouns]. *Zhongguo Yuwen [Studies of the Chinese Language]* 1964: 279, 263.