Epistemic Indefinites and Reportative Indefinites in Cantonese
Tommy Tsz-Ming Lee1 and Hok-Yuen Wong2

1University of Southern California; 2The Chinese University of Hong Kong

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 indeed conventional implicatures. We also reveal their various 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 [Cantonese]
Aaming seem eat-PERF what
Int.: ‘Aaming seems to have eaten something.’

b. Aming haoxiang chi-le shenme [Mandarin]
Aming seem eat-PERF what

Instead of bare wh-expressions, Cantonese adopts a numeral expression (i.e. (one)-CL NP) to express (4a).
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 ‘not-know-*wh*’ sequence
Aaming houci sik-zo  *m-zi matje*
Aaming seem  eat-PERF  *M-ZI what*
‘Aaming seems to have eaten something (which I don’t know).’

(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 (mentioned to the speaker before).’

Additional data suggest that these sequences are semantically similar to indefinites. They can establish dependency with regard to pronominal coreference and binding.

(8) Pronominal coreference
Aaming giu  {
   *m-zi bingo / bingo-bingo*;
   zoeng keoi-gaa ce zaa-gwolei wo5
}  zoeng keoi-gaa ce zaa-gwolei wo5
Aaming ask  *M-ZI who who-who*  zoeng keoi-gaa ce zaa-gwolei wo5
   *ZOENG he-CL*  car drive-here  *SP*
‘Aaming asks [someone (who I don’t know/ mentioned to the speaker before)]i to drive hisi car here.’

(9) Reflexive binding
Aaming  waa baan-japmin  {
   *m-zi bingo / bingo-bingo*;
   hou zungji caan
}  hou zungji caan
Aaming say  *class-inside M-ZI who who-who*  zoeng keoi-gaa ce zaa-gwolei wo5
   *very like*  praise zigeiwo5
   *SP*  who-who
   *himself*  *SP*
‘Aaming said that [someone (who I don’t know/ mentioned to the speaker before)]j likes to praise himselfi very much.’

However, these usages convey something more than a canonical indefinite expression. For example, (6) conveys the speaker’s ignorance on what Aaming ate, 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 the usages 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 the various distributional similarities of EIs and RIs. We give our syntactic and semantic
proposal for EIs and RIs in §4, followed by some discussions on the proposal in §5. We conclude the paper in §6.

<table>
<thead>
<tr>
<th>Table 1 - Summary of the claims</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Form</strong></td>
</tr>
<tr>
<td>----------</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Meaning</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>External syntax</strong> (§3.1)</td>
</tr>
<tr>
<td><strong>Internal syntax</strong> (§3.2 &amp; 3.3)</td>
</tr>
<tr>
<td><strong>Analysis</strong></td>
</tr>
<tr>
<td><strong>Proposed structure</strong> (§4.1)</td>
</tr>
<tr>
<td><strong>Proposed semantics</strong> (§4.2)</td>
</tr>
</tbody>
</table>

2. **Indefinites with additional implications**

2.1. *M-zi-wh* as Epistemic Indefinites (EIs)

*Claim*: The *m-zi-wh* sequence (as in (6)) is used as an EI.

*Note*: 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.

*Diagnostic tests: “Namely”-test / “guess-what” test*

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 fail 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 (which I don’t know).’
b. # …zikhai Hunglaumung
    that’s Dream.of.the.red.chamber
    ‘…That’s 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 an indefinite NP, 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.’

A note on morphology:
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.

2.2. Wh-doublets as Reportative Indefinites (RIs)

Claim: Wh-doublets are used as RIs in Cantonese.¹

Note: 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). RIs always appear in reportative contexts. The absence of such contexts results in marginality.

Diagnostic tests:
(10) Verbs of sayings: waa ‘say’
    ??(Aaming waa) keoi maai-zo bin-bun-bin-bun syu
    Aaming say he buy-PERF which-CL-which-CL book
    ‘Aaming said he bought some book (mentioned to the speaker before).’

(11) Hearsay evidential particle wo5
    Aaming maai-zo bin-bun-bin-bun syu ??(wo5)
    Aaming buy-PERF which-CL-which-CL book SP-hearsay
    ‘(It was said that) Aaming wants to buy some book (mentioned to the speaker before).’

Unsurprisingly, an out-of-the-blue context renders the use of RIs infelicitous in (12).

¹ 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.
(12) [The speaker began telling a story.]
   # Cungcin, Aaming maai-zo bin-bun-bin-bun syu
   once Aaming buy-PERF which-CL-which-CL book
   ‘Once, Aaming bought some book (mentioned to the speaker before).’

The reportative implication is removed once the RI is replaced by an indefinite NP. Compare (10) and (13). In (10), the speaker must have heard about the book titles, for example, from Aaming in some previous conversation, while in (13) it need not be the case.

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

2.3. Conventional implicatures (CIs) in EIs and RIs
2.3.1. ≠ Conversational implicatures
In the literature pertinent to EIs, the ignorance component of EIs can be attributed to conversational implicature based on Gricean 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 suggest that the ignorance component in EIs cannot be cancelled or be reinforced without redundancy, as suggested by the infelicity of (14b) and (14c) as continuations of (14a).

(14) a. Aaming gin-dou m-zi bingo
   Aaming see-ACH M-ZI who
   ‘Aaming saw some person (who I don’t know).’

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. (15) asserts the existence of an assignment that Aaming didn’t submit, one that the speaker cannot identify (i.e. the ignorance component remains). (15) 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).

(15) Aaming mou-gaau m-zi bin-fan gungfo
   Aaming not-submit M-ZI which-CL assignment
a. ‘Aaming didn’t submit some assignment (which I don’t know).’
b. # ‘Aaming didn’t submit any assignment.’

On a similar line, EIs in Cantonese are free from the anti-singleton constraint.

(16) Aaming hai syudim maai-zo m-zi bin-bun zeoi-gwai ge syu  
Aaming at bookstore buy-PERF M-ZI which-CL most-expensive NOM book  
‘Aaming bought some most expensive book (which I don’t know) at the bookstore.’

The ignorance component comes from the speaker’s failure to identify the book in a relevant way, say, by its name. The speaker identified it only by its description (i.e. the most expensive book). The anti-singleton constraint is considered crucial for a conversational implicature account (see discussion in Kratzer and Shimoyama, 2002; Alonso-Ovalle and Menédez-Benito, 2010; Alonso-Ovalle and Shimoyama, 2014).

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 dou mou-gong-gwo  
but he actually what book.title 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 syu wo5  
he not.have buy which-CL which-CL book SPhearsay  
a. ’He didn’t buy a book (mentioned to the speaker before).’
b. # ‘He didn’t buy any book.’

2.3.2. Conventional implicatures
We propose, instead, that both additional meaning associated with EIs and RIs are indeed conventional implicatures (hereafter ClIs) 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))
Potts (2005) adds that, following Grice (1975), CIs are generally commitments made by the speaker of the utterance by virtue of the meaning of the chosen words. Moreover, at-issue content in (19c) can be construed as ‘what is said’ (Grice 1975) or ‘regular asserted content’ (Potts 2005).

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...” suggests that only the CI content, but not the at-issue content, is going to be disputed. 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 (which I don’t know).’
   B: Hai aa3, batgwo nei zi keoi sik-zo matje gaak3
      yes SP but you know he 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 (mentioned to the speaker before).’
   B: Hai aa3, batgwo mou-juan gong-gwo hai matje jesik bo3
      yes SP but no-person say-EXP be what food SP
      ‘Yes, but no one mentioned the food he has eaten.’

Combining the claims in §2.3.1 and §2.3.2, 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. quelqu‘ (Jayez and Tovena, 2008, 2013), and RIs in Bulgarian (Koev, 2017)).

---

2 We do not further distinguish between CI and presupposition, for its irrelevance here. But see discussion in Koev (2017) and Potts (2005).
3. Distributional similarities between EIs and RIs

3.1. External syntax and scope behaviors

As mentioned in §1, both EIs and RIs are indefinites. Their distribution pattern with the 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/ bin-fan-bin-fan gungfo / jat-fan gungfo}, have M-ZI which-CL assignment which-CL-which-CL assignment one-CL assignment Aaming m-geidak zou wo5 Aaming not-remember do SP ‘There is some assignment that Aaming forgot to do {I don’t know which/ the assignment is mentioned to the speaker before/ there is one assignment}.’

However, while the Num-Cl-NP indefinites cannot appear in subject/ topic position (Li and Thompson, 1981), both EIs and RIs are not distributionally restricted in this aspect.

(23) *{M-zi bin-fan gungfo / bin-fan-bin-fan gungfo} Aaming mou M-ZI which-CL assignment which-CL-which-CL assignment Aaming not.have zou wo5 do SP ‘As for some assignment, Aaming didn’t do (which I don’t know/ the assignment is mentioned to speaker before)’.

In terms of interpretation, EIs and RIs usually take wide scope.\(^3\) We have already seen examples in (15) and (18) from §2.3, where both of them (inversely) scope over negation. (24) and (25) provide two more examples supporting their wide scope behaviors.

(24) **Inverse scope: EI/RI > every**

Mui-go jan dou tai-zo {m-zi bin-bun syu / bin-bun-bin-bun every-CL person all read-PERF M-ZI which-CL book which-CL-which-CL syu} wo5 book SP ‘Everyone read some book (which I don’t know/ mentioned to the speaker before).’

(25) **Island insensitivity: EI/RI > the majority**

Mui-go jan dou zidou complex-NP Aaming soeng maai {m-zi bin-zek every-CL person all know Aaming want buy M-ZI which-CL gupiu / bin-zek-bin-zek gupiu} ge beimat] wo5 stock which-CL-which-CL stock GE secret SP

\(^3\) Wide scope EIs have also been reported in Tiwa (Dawson 2018).
‘Everyone knows the secret that Aaming wants to buy some stock (which I don’t know/mentioned to the speaker before).’

Two differences with the one-Cl-NP indefinites should be noted. First, in (24), replacing EIs and RIs with one-Cl-NP indefinites give a strongly preferred surface scope reading. Second, in both cases, the sentences are unambiguous, i.e. there is no scope interaction. Thus, EIs and RIs here have a fixed wide scope reading, similar to a definite or referring expression.4

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. Table 2 summarizes various interpretations of *wh*-expressions when they occur in different environments.

(26) Aaming **m-zi dimjoeng** jeng-zo coeng beicoi
Aaming **M-ZI how** win-PERF CL competition
‘Aaming won the competition in some way (which I don’t know).’

(27) Aaming **dimjoeng-dimjoeng** jeng-zo coeng beicoi wo5
Aaming **how-how** win-PERF CL competition SP
‘Aaming won the competition in some way (mentioned to the speaker before).’

3.3. Strict adjacency

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

(28) a. *Aaming sik-zo **m-zi** [Num-CL saam-go] matje
Aaming eat-PERF **M-ZI** three-CL *what*
Int.: ‘Aaming ate some three things (which I don’t know).’

---

4 The fixed wide scope property may be attributed to the presence of *bin* ‘which’, which is linked to specificity. We set aside the interaction of scope behaviors and specificity. The main point here is that wide scope reading of EIs and RIs is available.
b. *Aaming sik-zo m-zi [Possessor Aafan ge] matje
   Aaming eat-PERF M-ZI Aafan POSS what
   Int.: ‘Aaming ate some Aafan’s thing (which I don’t know).’

A complication here is that there are cases where m-zi can occur sentence-initially (29a) or preverbally (29b), giving almost identical meaning. However, m-zi here is arguably different from its use in EI (see Appendix 1).

(29) 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’ (30). Also, no wh-expression in RIs can be extracted (31).

(30) Aaming hai bindou (*hai) bindou maai-zo syu wo5
   Aaming at where at where buy-PERF book SP
   ‘Aaming bought books at somewhere (mentioned to the speaker before).’
(31) *Geisi1, Aaming waa keoi t1 geisi maai-zo syu
    when Aaming say he t1 when buy-PERF book
    Int.: ‘Aaming said he bought books at some time (mentioned to the speaker before).’

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 ‘and’. However, this plurality requirement is found to be void in Cantonese, as singular reading is available (Wong, 2018). 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 \textit{matje} ‘what’ is two-layered:

(36) a. Normal semantics: \([matje] = \text{undefined}\)

b. Focus semantics: \([matje_F] = \{x: x \notin \text{human}\}\)

We propose that the internal structure of both EIs and RIs are the same, schematized as (37), where \textit{m-zi} in EI and the first \textit{wh}-expression in RI, abbreviated as \textit{WH}, head a functional position, say Y, taking a \textit{wh}-expression as their complement. \textit{M-zi} and \textit{WH} project a YP. Note that the category of Y depends on the category of their complement which can be a \textit{wh}-nominal or a \textit{wh}-adverbial (see also discussion in §5.1).

(37) The syntactic structure of EIs and RIs

\[
[YP \ m-zi_{i} / WH_{i} [XP \ \text{wh(-expression)} ]] 
\]

For the semantics of \textit{m-zi} and \textit{WH}, we propose that they are \textit{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 \textit{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 (38). Following Beck (2006), we assume a syncategorematic rule for these focus-sensitive operators as in (39).

(38) Semantics of \textit{m-zi} and \textit{WH} (first version, their implications aside)

\[
[m-zi_{i} / WH_{i}]_{g} = g(i) \in D_{cf} 
\]

(39) A special compositional rule for EIs and RIs

\[
[m-zi_{i} / WH_{i} XP] = [m-zi_{i} / WH_{i}] ([XP]^F) 
\]

For RIs, we suggest that the phonetic features of \textit{WH} is acquired during derivation. In particular, it copies whatever phonetic features of its complement at the PF component (resulting in reduplication, see Appendix 2 for complications in the formation of RIs).

As an illustration, the structure of (6) and (7) can be represented by (40), setting aside the semantics of the adverb \textit{houci} ‘seem’ and the particle \textit{wo5}. Given the wide scope interpretation of EIs and RIs (§3.1), we assume that the existential closure of choice function variables takes widest scope.

(40) \([TP2 \ \exists f [TP1 Aaming [VP sik-zo [DP2 m-zi_{1} / WH_{1} [DP1 matje]]]]]]\]

Let us first compute the semantics for the EI and RI in (40):
(41) $\llbracket DP2 \rrbracket = \text{(by (39))}$
\[
\llbracket m\text{-zi}_1 / \text{WH}_1 \rrbracket (\llbracket \text{matjier} \rrbracket^F) = \text{(by (36), (38))}
\]
\[
g(1) \in D_{ef}(\{x: x \notin \text{human}\}) =
\]
\[
f(\{\text{the apple, the cake, the cookies, …}\})
\]

Assume the semantics of sik-zo ‘ate’ to be (42):

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

The semantics for (40) can be computed as follows:

(43) $\llbracket TP2 \rrbracket = \exists f[\llbracket \text{VP} \rrbracket(\text{Aaming})] =$
\[
\exists f[\llbracket \text{sik-zo} \rrbracket(\llbracket DP2 \rrbracket)(\text{Aaming})] = \text{(by (41))}
\]
\[
\exists f[\llbracket \text{sik-zo} \rrbracket(f(\{\text{the apple, the cake, the cookies, …}\})(\text{Aaming})) = \text{(by (42))}
\]
\[
\exists f[\text{Aaming has eaten } f(\{\text{the apple, the cake, the cookies, …}\})]
\]
(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 conventional implicatures in EIs and RIs

Recall that the ignorance component in EIs and the reportative implication in RIs are conventional implicatures, rather than conversation 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:

(44) $\llbracket \cdot \rrbracket = < \{\text{at-issue content}\}, \{\text{CI}\}>$

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 (38) and the multidimensional analysis of CI, we propose the following semantics for EIs (45) and RIs (46), which maintain the parallelism as existential indefinites and capture their difference with regard to CIs:

(45) Semantics for $m\text{-zi}$ (final version)
\[
\llbracket m\text{-zi} \rrbracket^g = < g(i) \in D_{ef}, \text{fail.to.\underline{identify}(the\ speaker, g(i))} >
\]
Semantics for \( WH \) (final version)

\[
\llbracket WH \rrbracket^g = < g(i) \in D_{cf}, \exists s. \textbf{mentioned.to.the.speaker}(s, g(i)) > \text{ (where } s \text{ is some speaker)}
\]

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

\[
\llbracket (6) \rrbracket^g = < \exists f [\text{Aaming seems to have eaten } f(\{\text{the apple, the cake, the cookies, \ldots}\})], \text{ \textbf{fail.to.identify}(the \textbf{speaker}, f) } >
\]

\[
\llbracket (7) \rrbracket^g = < \exists f [\text{Aaming seems to have eaten } f(\{\text{the apple, the cake, the cookies, \ldots}\})], \exists s. \textbf{mentioned.to.the.speaker}(s, f) >
\]

5. Discussions and implications

The proposal in §4 captures the core meaning of EIs and RIs as (wide scope) indefinites (§3.1) by adopting a choice function analysis. It also explains why \( wh \)-adverbials are possible arguments for \( m-zi_i \) and \( wh_i \) (§3.2) as they denotes sets of alternatives in their focus semantics. The strict adjacency (§3.3) can be explained by the conventional close relation between a head and its complement. The conventional implicatures observed in §2 is accommodated in a two-dimensional semantics, following Potts (2005) and Koev (2017). This section discusses some related issues to the proposal. The examples below only include EIs but the same applies to RIs as well.

5.1. The syntactic status of the choice function variables

We remain vague in the syntactic category of \( m-zi_i / WH_i \). As we put it, they head a functional projection YP, taking a \( wh \)-expression as their complement. The category of Y depends on the category of their complement:

\[
\text{(49) a. } [\text{DP } \textbf{m-zi } [\text{DP bin-jat-bun syu}]] \\
\quad \text{M-ZI \ which-one-CL book} \\
\text{ b. } [\text{DP go } [\text{NumP jat } [\text{ClP bun } [\text{NP } \textbf{m-zi } [\text{NP matje syu}]]]]] \\
\quad \text{that one CL M-ZI \ what book} \\
\text{ c. } [\text{AdvP } \textbf{m-zi } [\text{AdvP dimjoeng }]] \\
\quad \text{M-ZI \ how}
\]

An alternative approach to capture the pattern in (49) is adjunction (instead of complementation). The distinction between an adjunction account and a complementation account here is reminiscent of the distinction in Q-adjunction and Q-projection language, as discussed extensively in Cable (2010). The distinction manifest in whether the Q can be separated from the \( wh \)-expression. Only Q-adjunction languages (e.g. Japanese), but not Q-projection ones (e.g. Tlingit and Sinhala), allow separation. As we have seen in §3.3, strict adjacency is observed between \( m-zi_i / WH_i \) and the \( wh \)-expression in EIs and RIs. We maintain the head status of \( m-zi_i \) and \( WH_i \).
5.2. Bare wh-expressions in Cantonese
Under the current proposal, we now have an account for why (4a), repeated as (50), is disallowed for an existential reading. (50) only allows an interrogative reading. Given that wh-expressions lack normal semantics and only come with focus semantics, the absence of a focus-sensitive operator leads to failure in interpretation of the wh-expression. On the other hand, the presence of an interrogative operator, presumably focus-sensitive, rescue the interpretation, but lead to an interrogative reading instead of an existential one.

(50) *Aaming houci sik-zo matje
   Aaming seem eat-PERF what
   Int.: ’Aaming seems to have eaten something.’

Note that the presence of a VP-level existential closure (Diesing 1992) is irrelevant here, since wh-expressions cannot be interpreted properly without a focus-sensitive operator instead of an existential closure. Similarly, the presence of an existential verb jau does not help either:

(51) *Jau matje Aaming sik-zo
    have what Aaming eat-PERF
   ‘Aaming have eaten something.’

In other words, bare wh-expressions in Cantonese cannot be used as indefinites because they lack a proper binder (i.e. a focus-sensitive operator). We also correctly predict the lack of an interrogative reading in (52a) and a universal reading in (52b). Since m-zi is the closest focus-sensitive operator to the wh-expression (compared to the interrogative operator and the dou operator), it blocks the binding of other higher operators on the wh-expressions.

(52) a. M-zi bingo heoi
    m-zi who go
   ‘Someone will go.’ / *‘Who will go?’

b. M-zi bingo dou heoi
    m-zi who all go
   ‘Someone will also go.’ / *‘Everyone will go.’

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). We also

5 The same explanation does not apply to the Mandarin counterpart as in (4b). We leave the issue to future research.
reveal their various distributional similarities (§3). We proposed a unified syntactic and semantic analysis for both EIs and RIs in Cantonese (§4).

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 (51a), 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.

(51) a. Aaming jam-zo m-zi seoi-ding-caa (le1)
    Aaming drink-PERF m-zi water-or-tea SP
    ‘Aaming drank either water or tea (and I don’t know which).’

b. *Aaming jam-zo seoi-ding-caa-seoi-ding-caa wo5
    Aaming drink-PERF water-or-tea-water-or-tea SP
    Int.: ‘Aaming drank either water or tea (and this has been mentioned to the speaker before).’

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.
Appendix 1: Preverbal and postverbal $m$-zi

$M$-$zi$ can appear in different positions in a sentence: sentence-initial position as in (52a), post-subject position as in (52b) and the post-verbal position as in (52c).

(52) 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

$$\approx \text{‘Aaming ate something (which I don’t know).’}$$

Although the meaning difference is subtle, we argue that $m$-$zi$ in (52a) and (52b) (hereafter pre-verbal use of $m$-$zi$) are (matrix) predicates (cf. Yap & Chor 2014), whereas that in (52c) (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 (52a), where the subject ‘I’ is null and they derive (52b) from (52a) 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:

(53) a. (Ngo) $m$-$zi$ Aaming sik-zo matje
    I M-$ZI$ Aaming eat-PERF what
b. Aamingi ngo $m$-$zi$ {?t / $\overline{OK}$keoi} sik-zo matje
    Aaming I M-$ZI$ he eat-PERF what

The marginality of (53b) 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 (52b). Note that it is substantially better than (54), where post-verbal $m$-$zi$ cannot take any external argument.

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

Although they do not discuss the case of post-verbal $m$-$zi$, it is possible, following their line of reasoning, that grammaticalization takes place further, depriving $m$-$zi$ 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 le1. Witness the contrast in (55):

(55)
le₁ is compatible with both matrix questions and embedded questions, but not with declarative sentences.

(56) a. Bindou hai le₁?  
    where be SP  
    ‘How come?’  
    (M&Y 2011:427)  

b. Nei waa [keoi hai-m-hai zeonbou-zo le₁]?  
    you say he be-not-be improve-PERF SP  
    ‘Do you think he has improved?’  
    (M&Y 2011:410)  

c. *Keoi hai zeonbou-zo le₁.  
    he be improve-PERF SP  
    Int.: ‘He has improved.’

It is known that m-zi, as a predicate, can take interrogative complement, forming embedded question and their compatibility with le₁ in (55a) and (55b) follows. The unacceptability in (55c) can be explained by suggesting that no embedded question is formed. If (55c) does not form an embedded question, m-zi in (55c) is not taking an interrogative complement, different from m-zi in (55a) and (55b). A natural explanation is that m-zi in (55c) is not a predicate (but a focus operator which does not form embedded question).

The last argument concerns possible continuations of the sentences in (52). (57) is a felicitous continuation of (55a) and (55b), but not (55c).

(57) ... daan honang me-dou mou-sik  
    but possibly what-all not.have-eat  
    ‘…but, maybe, nothing has been eaten.’

Although the embedded questions in (55a) and (55b) presuppose Aaming ate something, it can be retreated by (57). However, the infelicity of (57) as a continuation of (55c) suggests that the presupposition cannot be retreated. Intuitively, the speaker is making a contradictory statement with both (55c) and (57). It is suggestive that the proposition that Aaming ate something is not presupposed, but asserted in (4c). M-zi in (55a) and (55b), 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 (55c) suggests m-zi does not serve as a matrix predicate.
Appendix 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:

(58)  
(a) *Keoi gin-dou bin-bin-go wo5  
‘Aaming said he saw someone (mentioned to the speaker before).’  
b. Keoi gin-dou bingo-bingo wo5  
c. Keoi gin-dou bingo-bingo-jan wo5  
d. ?Keoi gin-dou bingo-jan-bingo-jan wo5

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 (58a) verifies the phrasal requirement in RI reduplication (on the assumption that bin ‘which’ is a D head). This proposal also correctly predicts (58b), as the whole phrase bingo, presumably a DP, is copied.

However, (58d) 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 (58c) (where NP ellipsis occurs) and the marginal (58d) (where NP ellipsis is absent). The same applies to matje ‘what’:

(59)  
(a) Keoi maai-zo bun matje-matje wo5  
(phrase copying)  
‘He bought some book (mentioned to the speaker before).’  
b. Keoi maai-zo bun matje-matje-syu wo5  
(phrase copying & NP ellipsis)  
c. ?Keoi maai-zo bun matje-syu-matje-syu wo5  
(phrase copying)

As for Reportative Adverbials (RAs), if we assume *WH-adverbials are phrasal, then the pattern in (60) follows. PF Copying at non-minimal phrasal level is disallowed.

(60)  
(a) Keoi dimjoeng-dimjoeng caau dip coi wo5  
(phrase copying)  
‘He fried the vegetables in some way (mentioned to the speaker before).’  
b. *Keoi dimjoeng-caau-dimjoeng-caau dip coi wo5  
(non-phrase copying)  
c. *Kaa keoi dimjoeng-caau-dip-coi-dimjoeng-caau-caau dip coi wo5  
(non-minimal phrase copying)

Another type of RI variants is *WH-triplets (61). 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.

(61)  
(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:

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


