The Genesis of Indefinite Pronouns in Japanese and Korean

TOMOHIDE KINUHATA  
*Osaka University*

JOHN WHITMAN  
*Cornell University*


1. Introduction

Japanese (1a) and Korean (1b) exemplify a crosslinguistically well-attested pattern where indefinite pronouns are composed of an interrogative word and an interrogative particle.

(1)  a. John ga nani-ka o katta.  
     John NOM what-ka ACC bought.  
     ‘John bought something.’

   b. John i mues-in-ka lul sassta.  
     John NOM what-be-ka ACC bought.  
     ‘John bought something.’

In this paper, we present a new account of the genesis of the indefinite pattern: Indefinite pronouns originate from a biclausal structure in which the interrogative word is an element in a question adjoined to the main
(2) [[s ... what, ... ?] [s Mina bought e,]]

The source pattern in (2) explains the placement of the interrogative particle, the presence of the copula *i-* in Korean, and parallels with diachronic development of particles like *yara* in Japanese (cf. Kinuhata et al, 2009). It also provides an account for the genesis of indefinites in terms of categorical relabeling (Whitman 2001).

2. Previous Analyses

2.1. Typological Accounts: Haspelmath 1997

Haspelmath (1997) claims that the Korean wh-*in-ka* pattern in (1b) originates from what he calls ‘parametric concessive conditionals’ like (3a).

(3) a. You can take something, whatever it, may be. p. 136
   b. You can take whatever it may be.

The main reason for this claim appears to be that the wh-*in-ka* pattern contains a reflex of the copula *i-*. However in every other respect the Korean pattern fails to share the properties of the ‘it may be’ pattern as shown below:

(4) ‘it may be’ pattern (Haspelmath 1997: 136-7)   Korean wh-*in-ka*

| Contains focus particle ‘also, even’ | None |
| Uses subjunctive mood or conditional marker | None |
| Uses a pleonastic negator or ‘want’ | None |
| Meaning: free choice ‘any’ | Meaning: existential |

Concessive conditionals with focus markers (Korean -*na* ‘even’, Japanese-*mo* ‘even, also’) are appropriate sources for expressions like Korean wh-*i-tunci* ‘any’, wh-*i-na* ‘any at all’ (Lee, Chung and Nam 2000, Yoon 2004), and Japanese wh-*de-mo* ‘any at all,’ but not for wh-*in-ka* and wh-ka. The former receive the same free choice interpretation identified by Haspelmath as the ‘original meaning’ (1997: 139) of the ‘it may be’ pattern. Haspelmath is surely right that there is a correlation between concessive conditionals as marked by the morphosyntactic devices in (4) and derived free choice pronouns. But the Japanese and Korean existential expressions in (1) share neither this morphosyntactic marking nor the free choice inter-

---

1 A diachronic source similar to this is proposed for the Korean pattern in (1b) by Yoon (2004).
pretation. Therefore they must come from another source.

2.2. kakarimusubi Analysis of Japanese Indefinites

Specialists in the history of Japanese have claimed that the historical source of the wh-ka indefinite pronoun pattern is the kakarimusubi construction in premodern Japanese (Sakakura 1975, Ogawa 1976-7, Yamaguchi 1991, Watanabe 2002). In the kakarimusubi construction, the scope of a constituent marked with what has traditionally been called a kakari particle is demarcated by a particular form of the predicate (traditionally labeled musubi). The kakari particle ka marks questions and occurs in construction with the focus of the question, while the adnominal (rentaikei) form of the predicate marks the scope of question. Thus in the following example kaku-er-u, the adnominal form of the auxiliary -er-i affixed to the verb kakus-‘hide’ marks the scope of the question.

(5) Tukwi wo ... kumo ka kakus-eru (雲香隠流)
   moon ACC cloud ka hide-RES[adm]
   ‘Are the clouds hiding the moon?’ Man’yoshu 1079, 8th C.

When used with case particles and postpositions, ka must follow them, as shown in the following Early Middle Japanese examples:

(6) a. Nani wo kamo kofi to fa ifu
   what ACC ka love COMP TOP say
   ‘What does (each person) say to be love?’ Ise monogatari, 10th C.

b. [[Ikayau naru kokorozasi ara-mu fito] ni] ka
   what.kind be intention have-Guess person DAT ka
   afa-mu to oposu
   marry-Guess COMP think[adm]
   ‘A person who has what kind of intentions do you think you would like to marry?’ Taketori monogatari, 9th C.

If we assume that PPs and case-marked NPs in premodern Japanese are islands, then the generalization that ka and the adnominal ending is restricted by subjacency (Yanagida 1995, Whitman 1997) explains the position of this particle. If ka precedes the case marker or the postposition, it cannot be subjacent to the adnominal predicate form marking the scope of the question, due to the NP/PP island.

A consequence of this is that ka in kakarimusubi can be separated from the interrogative word. Thus ka is separated from the interrogative word by the accusative particle in (6a) and the intervening portion of the complex NP and the dative particle in (6b). In contrast, in modern Japanese wh-ka indefinites, case markers or postpositions cannot intervene between ka and the indefinite pronoun, as shown in (7).
THE GENESIS OF INDEFINITE PRONOUNS IN JAPANESE AND KOREAN / 91

(7) a. Taro wa {nani ka o/nani o ka} katta.
   Taro TOP what ka ACC/ what ACC bought
   ‘Taro bought something’

b. {Doko ka de/ doko de ka} ongaku o ensou si-teiru.
   where ka in/ where in ka music ACC play do-PROG
   ‘Music is played somewhere.’

Thus we see that the position of ka relative to the interrogative word differs in the premodern kakarimusubi construction and modern Japanese indefinites.2

This means that any analysis which posits the kakarimusubi construction as the source of wh-ka indefinite pronouns must account both for the change of sentence type (from interrogative to declarative) and the change of the position of ka with respect to the interrogative word. However, none of the analyses that have been proposed explicitly explains these changes.3

Even if the two changes above are explained theoretically, there remain two empirical problems for any account taking the kakarimusubi construction to be the source of wh-ka indefinites. First, there is approximately a 200-year-gap between the disappearance of the kakarimusubi construction and the emergence of indefinite pronouns with ka. We can still find examples of kakarimusubi with ka in the Amakusaban Heike Monogatari, which is published in 1593, but we find none in the Toraakirabon Kyôgensyû (transcribed in 1642) except for those in poems. However, according to Ki-nuhata & Iwata (2010), it is reasonable to think that the indefinites with ka first emerge around the beginning of the 19th century. Second, a kakarimusubi-based account can at most explain the indefinite pronoun pattern with ka of Japanese. But the wh-ka pattern is just one of a set of similar constructions in Japanese and Korean containing a wh-type pronoun. None of these other constructions can be derived from the kakarimusubi construction. Korean, as is well known, has never attested this pattern, despite the evident parallelism between (1a) and (1b). Within Japanese as well, there attested indefinite pronouns without ka; i.e. indefinites consisting of ‘interrogative word + yara’ as in (8) below.

2 Some postpositions seem to allow ka to precede or follow it. See footnote 7.
3 For instance, Sakakura (1975) states that the weakening of the kakari function of ka caused the change of the property of ka. [係助動詞カ・ソが次第に係り機能を弱くして、ついに、係体助詞または接尾語としての性格を持つに至った (p. 252)] However, it is not clear to us why the ‘weakening of kakari function’ forces ka to precede the case markers as a suffix.
(8) **Nani-yara-ga mune-ni-fa ari-te**
what-yara-NOM heart-DAT-TOP exist-CONJ

‘I have something to say (, but I cannot tell).’ *Udukimomizi*, 1706

However, *yara* has never been used as a *kakari* particle. Thus another explanation is necessary for the origin of the indefinite pronoun pattern.

3. An Adjoined Question Source for wh-Q Indefinites

We propose that indefinite pronouns originate from what we call matrix-adjoined wh-questions as shown in (9) below (see Lee 1997, Takamiya 2004, Kinuhata et al. 2009 for the same source for similar but different constructions).

(9) **Nani o omotta no ka, Taro wa pinku no syatu o katta.**
what ACC thought FN ka, Taro TOP pink GEN shirt ACC bought.

‘What did he think, Taro bought a pink shirt.’

In (9), the speaker wonders why Taro bought a pink shirt with the adjoined question.

Since the main clause of a matrix-adjoined wh-question must be an assertion (not a question, imperative, volitive, etc.) as exemplified in (10), let us label their LF *Asser* as in (11).4

(10) a. *Nani o omotta no ka, pink no shirt o kainasai.*
what ACC thought FN ka, pink GEN shirt ACC buy[IMP].
*‘What did he think, buy a pink shirt.’*

b. *Nani o omotta no ka, pink no shirt o kaou.*
what ACC thought FN ka, pink GEN shirt ACC buy[VOL].
*‘What did he think, I will buy a pink shirt.’*

(11) [[Ques Nani o omotta no ka][Asser Taro wa pink no shirt o katta]]

We propose that the change of wh-ka from questions to indefinites occurs when the adjoined phrase loses its clausal status. Let us illustrate this with the following example.

---

4 In Kinuhata (2007), the interrogative sentence in examples such as (9) is not structurally adjoined to the following assertion sentence. Rather he considers it as being related to the following sentence in discourse, which accounts for the differences between *kakarimusubi* and indefinites. Though this line of thinking is still viable, we adopt the view in this paper that the interrogative sentence is structurally related to the following sentence.
(12) \textbf{Itu} kara \textbf{ka}, kirui wo siti ni ma wo watasi
when from Q clothes ACC pawn to interval ACC crossing

i. ‘From when was it, I have been making ends meet pawning clothes.’
ii. ‘From some point, I have been making ends meet pawning clothes.’

\textit{Tennoamizima}, 1720

As made clear by the translations, (12) has two interpretations from the viewpoint of our analysis: (i) is a matrix-adjoined question interpretation and (ii) is an indefinite interpretation.

Thus (12) has two LFs as in (13):

(13) i. LF: [[\textit{Ques} itu kara [kirui wo siti ni ma wo watasi] ka] [\textit{Asser} kirui wo siti ni ma o watasi]]

ii. LF: [[\textit{Asser} [itu kara ka] [\textit{Asser} kirui wo siti ni ma wo watasi]]]

Since ‘itu kara ka’ is a question in its matrix-adjoined question interpretation, it has clausal status as in (13i); we assume that the unpronounced (that is, sluiced) part of the question is copied at LF.\(^5\) However, when ‘itu kara ka’ gets an indefinite interpretation, it does not have clausal status as in (13ii). The loss of clausal status in (13ii) is triggered by the reanalysis of the sentence type. In (13i), the sequence is composed of two sentences: one is a question and the other an assertion. But in the LF of (13ii), the whole sentence is reanalyzed as an assertion, which leads to the loss of clausal status of the adjoined phrase.

This reanalysis of the LF causes a concomitant change in the semantics. If we consider the semantics of wh-questions as a set of propositions obtained by substituting the part that corresponds to the wh phrase (Hamblin 1973), the loss of clausal status in (13) leaves the substitution as the semantics of ‘itu kara ka’ as below.\(^6\)


6 Though Hamblin (1973) does not take the semantics of the wh-phrase to be an existential quantifier, Karttunen (1977) does. We follow the semantics of Karttunen (1977) in this respect. Note, however, that wh-ka indefinites in Japanese do not necessarily obey the Novelty Condition (Heim 1982) as the following example shows.

Taro wa nanika o hirotta, sosite sono nanika o tabeta.
‘Taro picked up something and then ate that something.’

This means that we have to revise the semantics of wh-\textit{ka}. In using wh-\textit{ka}, we have the intuition that the speaker does not know what the object is. If this is correct, we have to specify the meaning of wh-\textit{ka} relative to the knowledge of the speaker rather than the discourse. In this sense, the term ‘indeterminate pronoun’ used by Kuroda (1965) for Japanese wh-words might be a more appropriate term for the wh-\textit{ka} pattern as well.
(14) a. $\langle \text{[[} \text{Ques, itu kara [kirui wo siti ni ma wo watasi ta] ka} \text{]} \text{[Asser kirui wo siti ni ma wo watasi ta]}} \rangle_{M^g}$
   $= \lambda p \exists r [p = \text{from } t, \text{I have made ends meet pawning clothes}] \&$
   [ \text{I have made ends meet pawning clothes}]

b. $\langle \text{[[} \text{Asser, itu kara ka} \text{]} \text{[Asser kirui wo siti ni ma wo watasi ta]}} \rangle_{M^g}$
   $= [ \exists r [\text{from } t, \text{I have made ends meet pawning clothes}]]$

However, examples like (12) cannot be reanalyzed as NPs, i.e. as indefinite pronouns, since wh-ka includes the postposition kara[from] between wh and ka; thus the matrix-adjoined question interpretation is more prominent. This is supported by the fact that that ‘itu kara ka’ can only be used with assertions.\(^7\)

\(\begin{align*}
(15) \ a. & \text{??Itu kara ka sigoto o hazimenasai.} \\
& \text{when from ka work ACC start(IMP).} \\
& \text{‘Please start your work from some time.’} \\
\quad & \text{b. ??Itu kara ka sigoto o hazimeyou.} \\
& \text{when from ka work ACC start(IMP)} \\
& \text{‘I will start my work from some time.’}
\end{align*}\)

Wh-ka in (12) occupies a non-argument position, referring to the starting time of the event. On the other hand, in data where wh-ka denotes an argument, it is more difficult to interpret it as an independent clause. Thus, although the following example theoretically has two interpretations, the indefinite interpretation is more salient, with an LF like (17b).

(16) Zisuke ga te wo tori nani ka te ni watasu.
   Zisuke GEN hand ACC take what ka hand DAT give
   i. ‘(Tuyu) takes the hand of Zisuke and, what is it, she gives him.’
   ii. ‘(Tuyu) takes the hand of Zisuke and gives him something.’

\textit{Teikohonana}, 1807

(17) a. LF: $\langle \text{[[Ques, nani [te ni watasi-ta] ka} \text{]} \text{[Asser e, te ni watasi-ta]]}\rangle^8$
   b. LF: $\langle \text{[Asser, nani ka] [Asser e, te ni watasi-ta]]} \text{ or}
   \text{[Asser, nani ka te ni watasi-ta]}

In the early 19\textsuperscript{th} century data (as in modern Japanese), it appears that when

---

\(^7\) Wh-ka with a postposition between the wh-item and ka primarily has a matrix-adjoined wh-question interpretation primarily. This explains the difference in acceptability of the following pair.

\(\langle\text{Doko ka kara/ } \text{’doko kara ka} \text{) tegami o dasoo.} \)
\(\text{where ka from where from ka letter ACC send[VOL]}\)
\(\text{‘I will send the letter from somewhere.’}\)

\(^8\) We assume that e in the second clause is dynamically bound by the ‘nani’ in the first clause.
wh-ka is interpreted as an argument, case markers have a tendency to be dropped. This may have facilitated the reanalysis of wh-ka as an NP through an intermediate stage when it was co-indexed with an empty category in the main clause argument position.

The fact that wh-ka is integrated into the main clause as an NP causes an effect on the sentence in which it is used. As is shown by the following examples, the type of the sentence which includes wh-ka is not restricted to assertions since it is not a matrix-adjoined wh-question.

\[(18)\]
\[\begin{align*}
\text{a. } & \text{Nanika motte mairi-masyo.} \\
& \text{something bring come-POL(VOL)} \\
& \text{'I will bring something.'} \\
& \text{Keiseikaiyûkagami, 1818} \\
\text{b. } & \text{Nanika motte ki-nasai} \\
& \text{something bring come-HON(IMP)} \\
& \text{'Please bring something'}
\end{align*}\]

In summary, our hypothesis explains the change in sentence type and the position of the interrogative particles. The sentence type changes from a question-assertion sequence through an assertion, to finally any sentence type. Wh-ka in the adjoined position coindexed with an empty category gives it the status of NP, thus it precedes case markers. In the next section, we will present empirical data from both Japanese and Korean supporting these claims.

4. History of Indefinites

4.1 Japanese

In Early Middle Japanese, interrogative words can only be used in questions unless they are not bound by mo: e.g. dare mo (whoever). However, in Middle Japanese, the indefinite use of interrogative words begins to appear. The form of indefinites in this period consists of interrogative words and yarau (or yaramu/yaran) as (19).9

\[(19)\] Nani yarau de kono tyuu wo mita zo.
\[\text{lit 'I saw this annotation in something.'} \quad \text{Shikishô, 1477}\]

Yarau later becomes yara, dropping the final u as in (8). Yarau is originally from the sequence ‘ya+ara+mu [kakari particle+copula verb+conjectural marker]’. Though it includes a kakari particle, yarau itself was never used

---

9 According to Konoshima (1966), indefinite use of wh-ze also appeared in Middle Japanese. However, because of the space constraints here, we leave a complete study of the genesis of this pattern for further research.
as in a *kakarimusubi*. Before the emergence of its indefinite use, it was used like a sentence final particle with the sentence including it functioning as a question.

(20) ‘Are fa ikanaaru zyauru nite masimasu yaran’ to that TOP what noble.person COPULA HON CJEC QUOTE ‘Who is that noble person?’ *K Heike*, 14th C.

As a byproduct of this construction, sentences ending with *yaran* could also be used as matrix-adjoined questions as (21).

(21) Yo-ni-fa ikanisite more-keru-yaramu, afareni yasasiki world-in-TOP how leak-PAST-CJEC sad lovely tamesi-ni-zo fitobito mausi-aferi-keru. instance-as-EMP people talk-HABITUAL-PAST ‘How did people come to know, they were talking about the story, seeing it as sad and lovely.’ *K Heike*, 14th C.

In the *Kakuichibon Heike Monogatari*, 13 examples out of 58 are used as matrix-adjoined wh-questions. Since all the other examples of *yaran* with wh phrases exhibit the sentence final pattern as in (20), the indefinite pattern in (19) must have come from the matrix-adjoined use unless it is derived from sentence final *yaran* directly.

We should also note here that all the examples of *yaran* are followed by assertive clauses. This supports the claim that indefinite *yaran* evolved from matrix-adjoined questions.

According to Kinuhata & Iwata (2010), indefinite pronouns with *ka* appeared around the 19th century in the history of Japanese. However, they found pre-19th century examples of wh-*ka* with an adverbial use. The following data is taken from Kinuhata & Iwata with some modifications.10

<table>
<thead>
<tr>
<th>TABLE wh-<em>ka</em> in Japanese</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>Adverbial</td>
</tr>
<tr>
<td>Argument</td>
</tr>
</tbody>
</table>

The examples classified as ‘adverbial’ include those referring to time like (12) and non-argument quantifiers such as (22a). Wh-*ka* in argument positions like (16) and those that are coreferential with the NP in argument position like (22b) are classified as ‘argument.’

The Genesis of Indefinite Pronouns in Japanese and Korean

(22) a. Ikudo ka kiitare domo
    ‘Though I heard some times,’ Sinzyûkasaneidutu, 1707
b. Kamiire yori nanika kaitamono, wo idasu
    ‘He take out something written from his wallet.’ Hokkatûzyo, 1794

Table 1 shows that the formation of indefinites with ka starts with the phrase in the adjoined position rather than in argument position. Since, as we saw in the previous section, wh-ka in the adjoined position has more properties in common with the matrix-adjoined wh-questions than that in argument positions, it is reasonable to conclude that the former emerges in the language earlier than the latter.

Positing matrix-adjoined wh-questions as the source for indefinite use of ka faces one problem: it is usually claimed that ka was not used in sentence final position for wh-questions in Middle Japanese (see Sakakura 1975 among others). However, in Early Modern Japanese, we can in fact find some examples in which ka is used in the sentence final position of wh-questions. The following is an example of a matrix-adjoined question in which the narrator wonders why the speaker mistook the person in question.

(23) Mago no koto wa kimotuka-zu rougan no nani grandchild GEN thing TOP notice -NEG presbyopia NOM what mite ka, ‘Mumu, madu syokunin ni niawa-nu, ano see ka mumuu above.all tradesman GEN suit -NEG that bintuki ga kiniira nu.... hair.style NOM like -NEG ‘Not noticing his grandchild, what do my aged eyes see, “Mumu, I don’t like his hair style, which above all is not suitable for a tradesman”.’

Since zo was still predominant in the sentence final position of wh-questions, this type of example using ka is not numerous in Early Modern Japanese. But examples like (23) indicate that our analysis has more empirical support than the kakarimusubi analysis.11, 12

---

11 Yukinori Takubo (p.c.) pointed out to us that even if ka is not used in the sentence final position of wh-questions, the use of ka in indirect questions might give the evidence that ka can be used with wh-phrases. According to Kinuhata & Iwata (2010), indirect wh-questions with ka appeared around 18th century, in which we can find direct wh-questions with ka like (23).

12 It is hard to determine what type of sentences are used with wh-ka indefinites, since Kinuhata & Iwata (2010) exclude from their data examples of questions and hortatives in order not
4.2 Korean

Korean has a wide range of indefinite pronoun patterns that originate from wh-interrogative clauses, reviewed in detail by Yoon (2004). Yoon argues that a subset of these patterns have undergone a reanalysis from clause to nominal category similar to the one we argued for in Japanese in the previous section. That includes the wh-in-ka ‘wh-COPULA-Q’ pattern exemplified in (1b) and its counterpart using the suspensive complementizer -ci, wh-in-ci ‘wh-COPULA-SUSP. Martin (1992) provides an example of the latter:

(24) Nwuku-n-ci wa-ss-ess-ta.
    who-COP-SUSP come-PAST-RPAST-DEC
    ‘Someone came (while you were away).’ (Martin 1992: 738)

It is clear that the wh-in-ka ‘wh-COPULA-Q’ pattern is a relatively recent innovation. We know this because through the 18th century, the question particle ka was used only in yes/no questions, while wh-questions appeared with the [+wh] Q particle kwo. The wh-in-ka pattern is thus no earlier than the loss of the specialized [+wh] Q particle kwo in Seoul Korean in approximately the 18th century.

Both the wh-in-ka ‘wh-COPULA-Q’ and wh-in-ci ‘wh-COPULA-SUSP’ patterns show some speaker variation as to degree of grammaticalization. Some speakers find examples where they are followed by case particles somewhat substandard. Yoon (2004: 24) observes that wh-in-ci is significantly less acceptable when followed by a case marker. Nevertheless it is possible to find such examples:

(25) Nwuku-n-ci ka wa-se, way o-n ke nya.
    who-COP-SUSP NOM come-and why come-ADN fact Q
    ‘Someone came; why was it that they came?’
    yuuhi.egloos.com/1697601

For the purposes of our argument in this paper, the crucial property of the wh-in-ka ‘wh-COPULA-Q’ and wh-in-ci ‘wh-COPULA-SUSP’ patterns is that they clearly show the clausal origins of these indefinite pronouns. Both patterns contain a wh-pronoun (nwukwu ‘who’ in the examples above), the adnominal form of the copula (the form selected before clause-final question particles), and an interrogative particle.

Yoon (2004) discusses the proposal of Chung (1996) that all Korean existential patterns of the form wh-COPULA-Q have an internal clausal structure synchronically in contemporary Korean. Yoon rejects this hy-
pothesis in the original form proposed by Chung. She shows that instances of the wh-COPULA-Q pattern like (1b) which allow case marking co-exist with instances that still involve an adjoined interrogative clause. This is parallel to what we saw with Japanese examples such as (12). As the reanalysis from adjoined question to indefinite pronoun is in progress, both interpretations co-exist.

5. Conclusion

In this paper we have presented an account of the diachronic source of indefinite expressions composed of an interrogative word and an interrogative particle. We have argued that such expressions originate from wh questions adjoined to assertions. We have shown that this account is superior to the source posited for such expressions by earlier researchers on Japanese, who have claimed that wh-ka indefinites originate from kakarimusubi pattern. Our account is also compatible with the analysis of Japanese indefinites with yara and Korean indefinites such as wh-in-ka, which also originate from clausal adjuncts adjoined to a main clause assertion (Kinuhata 2007, Yoon 2004). It is also compatible with the account of syntactic reanalysis proposed in Whitman (2001), where the first step in a syntactic analysis involves a change in category (relabeling) rather than a change in hierarchical structure. In the case of the present change, the first step is the relabeling of the adjoined question to a non-clausal category, after Sluicing of its clausal content.

References

Kinuhata, T. & M. Iwata. 2010. Meishikuichi no ka no rekishi: Sengen, futei yōhō o chūshinni [The history of ka in NP-positions: Focusing on disjunctive and Inde-