C-5

Split Questions and Ellipsis in Japanese

Keisuke Yoshimoto
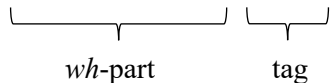
Ryukoku University

Abstract

In his paper on Spanish and English split questions, Arregi (2010) applies the analysis of short answers by Merchant (2004) and argues that the tag of split questions (e.g. *What did John plant, an oak?*) involve movement and deletion. In this presentation, I attempt to apply previous analyses of Japanese short answers (Saito 2004, Nishigauchi and Fujii 2006) to the split questions in Japanese (e.g. *Anata-wa nani-o tabemasita ka? Ringo(-o) hutatu desu ka?* 'What did you eat, two apples?'). It is argued that the underlining structure of the second part of split questions is a cleft sentence with the presuppositional CP subject deleted by argument deletion.

1 Introduction

Since the seminal work by Merchant (2001) on sluicing, a lot of attention has been devoted to the hidden structure of syntax. Split questions are one such structure that, according to Arregi (2010), involves movement and deletion along the same line with the analysis of short (fragment) answers by Merchant (2004). As shown in (1), split questions consist of a *wh*-part and a following tag.

(1)   What did John plant, an oak?

        wh-part   tag

The sentence in (1) is in fact bi-clausal and the tag is derived from a full clause. The underlying structure of the tag in (1) is (2) in which focus movement has taken place to the front of the clause and the remnant has been deleted (Arregi 2010)

(2) $[_{CP}$ $oak_{j\,F}$ $C_{F,\,E,\,Q}$ $[_{TP}$ ~~John planted t_j~~ $]]$

C has a Focus feature, Q feature and an E(llipsis) feature (Merchant 2001), and a tag has moved to CP Spec to check the Focus feature. Subsequently, the complement of C carrying an E feature is deleted, leaving behind the tag on the surface.

Inspired by Arregi (2010), I intend to investigate the derivation of Japanese split questions. Unlike English counterpart, Japanese split questions apparently come in two clauses; the first clause is a *wh*-question and the second a copular sentence as in (3). This can be expressed in more informal way without the politeness marker *-mas-* in the first part and the copular in the second part as in (4). The sentence in (4) is pronounced with the rising intonation.
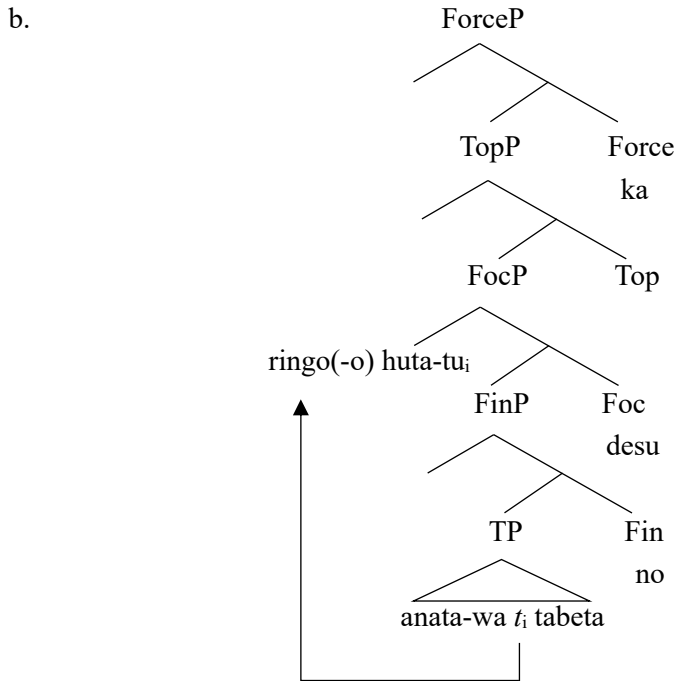
(3)   Anata-wa nani-o   tabemasita ka? Ringo(-o)   huta-tu desu ka?
       You-TOP   what-ACC ate   Q   apple(-ACC) two-CL COP Q
       'What did you eat, two apples?"
(4)   Kimi-wa nani-o   tabeta   no?   Ringo(-o)   huta-tu?
       You-TOP   what-ACC ate   Q   apple(-ACC) two-CL
       'What did you eat, two apples?'

Although the second part of the split question in (3) appears to be a bare copular sentence, I suggest that its underlying structure is a cleft sentence with the presuppositional CP subject deleted by argument deletion. Following Hiraiwa and Ishihara (2012), I hereby assume that Japanese cleft sentences are derived from in-situ focus constructions (*no da* constructions) through focus movement and subsequent topic movement. Thus, the second part of (3) begins with the in-situ focus construction in (5).

(5)   [ForceP [TopP [FocP [FinP [TP anata-wa   ringo(-o)   huta-tu   tabeta] no]   desu]] ka]?

                          You-TOP   apple(-ACC) two-CL   ate   no   COP   Q

    'Is it two apples that you ate?'

From (5), *ringo(-o) huta-tu* 'two apples' is moved to the FocP Spec position by focus movement, and we get (6a). The structure of (6a) is roughly diagramed in (6b).

(6) a. [ForceP [TopP [FocP ringo(-o)   huta-tu$_i$ [FinP [TP anata-wa t_i tabeta] no] desu]] ka]?

                     apple(-ACC) two-CL   you-TOP   ate   no   COP   Q

    'Two apples, did you eat?'

   b.

```
                        ForceP
                       /      \
                     TopP      Force
                    /    \       ka
                  FocP    Top
                 /    \
   ringo(-o) huta-tu_i  \
                  FinP    Foc
                 /   \    desu
               TP     Fin
              /  \     no
    anata-wa t_i tabeta
```

Subsequently, the remnant FinP is moved to TopP Spec by topic movement, resulting in a cleft sentence in (7a), whose tree diagram is shown in (7b).

(7) a. [ForceP ~~[TopP [anata-ga t_i tabeta no wa]$_j$~~ [FocP ringo(-o)   huta-tu$_i$ t_j desu]] ka]?

                    you-NOM   ate   C TOP   apple(-ACC)   two-CL   COP Q

    'Is it two apples that you ate?'

b.

ForceP

TopP   Force
           ka

~~anata-ga *t*ᵢ tabeta no-wa~~   FocP   Top

ringo(-o) huta-tuᵢ

FinP   Foc
       desu
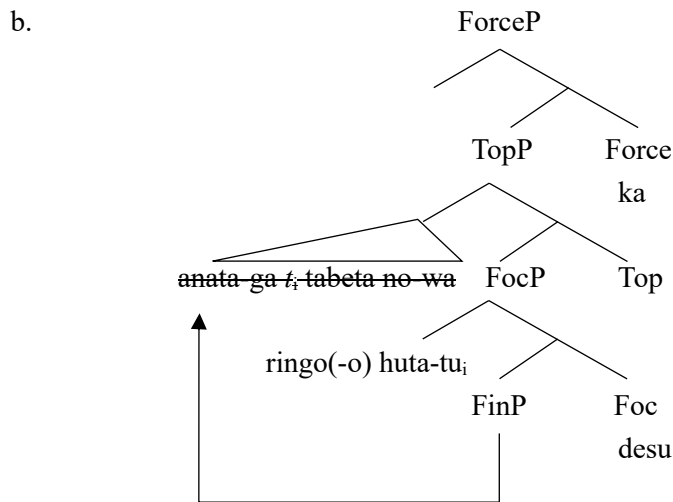
Here, I propose that the presuppositional CP subject of the cleft sentence (e.g. anata-ga t_i tabeta no-wa 'what you ate') is deleted by argument deletion. That argument deletion of subjects is possible in Korean and Japanese is argued for on independent grounds (Kim 1999, Oku 1998). Consequently, we get the second part of the split questions in (3).

This paper is organized as follows. In the next section, I present various connectivity phenomena that support the proposed analysis. In section 3, I discuss several approaches to island insensitivity of split questions.

2   Connectivity

In the previous section I presented an analysis of Japanese split questions according to which their second part is derived from cleft sentences through subject CP deletion. In this section, I adduce evidence in support of the underlying cleft structure. Note that these diagnostic sentences are the ones Nishigauchi and Fujii (2006) used for their analysis of Japanese short answers, and I applied them to split questions.

First, a case assigned to the NP in the second part matches the one assigned to the *wh*-phrase in the first part. In (3), the accusative case *-o* is attached to the NP in the second part, which matched that of the *wh*-phrase, *nani-o*, in the first part. If this were, for instance, the nominative case *-ga* or the dative case *-ni*, the sentence is not acceptable.

(8)   Anata-wa   nani-o   tabemasita ka?   Ringo(-o/*ga/*ni)   huta-tu   desu   ka?
        You-TOP   what-ACC   ate   Q   apple(-ACC/*NOM/*DAT)   two-CL   COP   Q
        'What did you eat, two apples?"

If the second part were a bare copular sentence, this case matching would remain a mystery. Furthermore, an analysis of cleft sentences in terms of base-generation (Kuzu 2005) would require an independent mechanism to assure the case matching. The proposed derivation of cleft sentences by Hiraiwa and Ishihara (2012) nicely captures this because the NP in the second part is assigned an accusative case by the verb *tabeta* 'ate' prior to focus movement.

Second, binding also reveals the concealed cleft structure. Let us see the binding of a reflexive pronoun *zibun-zisin* 'self'.

(9)   Ken-wa   dare-o   semeta   no desu ka?   Zibun-zisin   desu   ka?

Ken-TOP   who-ACC   blamed   C   COP   Q   self   COP   Q

'Who did Ken$_i$ blame, himself$_i$?'

In (9), the reflexive pronoun *zibun-zisin* 'self' is bound by *Ken*. This indicates that the second part consists of a full clause in which the reflexive pronoun is locally bound by the subject. On the other hand, a bare copular analysis cannot account for this fact because the only available antecedent for *zibun-zisin* for that analysis would be *Ken* in the first part, which is too far to bind it due to Condition A of the binding theory.

    Third, bound pronouns also suggest the presence of the hidden subject in the second part.

(10) a. Kaku sya$_i$-wa   dare-o   syootaisimasu ka? Soko$_i$-no kabunusi-o   hutari-zutu
       Each company-TOP who-ACC invite   Q   It-GEN   stockholder-ACC 2-CL-each
       desu ka?
       COP   Q
       'Who will each company$_i$ invite, their$_i$ two stockholders each?'

     b. Minna-wa   dare-o   paatii-ni   syootaisimasu ka?   Zibun(zisin)-no tomodati-o
       Everyone-TOP who-ACC party-DAT   invite   Q   Self-GEN   friend-ACC
       hutari-zutu   desu   ka?
       2-CL-each   COP   Q
       'Who will everyone$_i$ invite to the party, their$_i$ two friends each?'

In (10a), *soko* 'it' in the second part is interpreted as bound by *kaku sya* 'each company' in the first part. That is, the speaker asks whether companies will invite two stockholders each. Likewise, (10b) has a similar interpretation in which *zibun(zisin)* 'self' in the second part is bound by *minna* 'everyone' in the first part. This can be accounted for in the proposed analysis because the bound pronoun and its antecedent belong to the same clause (Binding Domain) and the former is c-commanded and bound by the latter in the base position prior to focus movement.

    Fourth, scope interaction confirms the structure I am proposing. In this light, let us see (11).

(11)   Minna-wa   samba-o   nankyoku   utaimasita   ka?
        Everyone-TOP   samba-ACC how.many.song   sang   Q
        Sankyoku   desu   ka?
        3-songs   COP   Q
        'How many sambas did everyone sing, three sambas?'

The second part of (11) has ambiguous interpretations with respect to scope. In one interpretation, everyone sang the same three sambas together in which 3 has a wider scope than the universal quantifier (3>everyone). In the other interpretation, everyone sings three sambas each (not necessarily the same sambas) in which the universal quantifier takes a wider scope than 3 (everyone>3). It is important to note that the same scopal ambiguity can be attested in a cleft sentence in (12).

(12)   [Minna-ga   utatta   no]-wa   samba-o   sankyoku   desu   ka?
        Everyone-NOM sang   C-TOP   samba-ACC   3-songs   COP   Q
        'Was it three sambas that everyone sang?'

(12) has both readings: everyone sang the same three sambas together (3>everyone) and each one of them sang three sambas each (everyone>3). This strongly supports the current idea that the second part in split questions is derived from a cleft sentence. On the other hand, it would be difficult for a bare copular analysis to account for why such scope ambiguity obtains in split questions.

Lastly, let us see 'vehicle change' of Fiengo and May (1994). 'Vehicle change' is a rule that transforms a name in an elided structure into a pronoun, so a Condition C violation is avoided. Merchant (2004) applies this to account for the lack of a Condition C violation in short answers in (13Ba).

(13) A: Who did you tell *t* about Bill$_i$'s raise?
      B: a. [$_{CP}$ Him$_i$ [I told t_i about his$_i$ name]].
          b. * I told him$_j$ about Bill$_i$'s raise.

(13Bb) is ruled out due to the Condition C violation, but that is salvaged in (13Ba) when ellipsis has taken place and changed the name into a pronoun. That the same applies to Japanese split questions can be seen in (14).

(14) Dare-ni   Ken$_i$-no   mozi-ga   kaidoku   dekiru no   desu   ka?
      Who-DAT   Ken-GEN   handwriting-NOM   decipher   can   C   COP   Q
      'Who can decipher Ken$_i$'s handwriting?'
      Kare$_i$-ni   dake   desu   ka?
      He-DAT   only   COP   Q
      'Only he$_i$ can do it.'

In (14), the pronoun *kare* 'he' in the second part is construed as *Ken* in the first part. If the second part contained a concealed structure in which *kare* 'he' c-commands *Ken*, it would violate Condition C. However, the violation is obviated due to 'vehicle change'.

So far, I have shown evidence that the second part of Japanese split questions involves ellipsis, which deletes the presuppositional CP subject of a cleft sentence. A natural question that arises at this point is regarding evidence for movement. In this regard, let us see the possibility of binding intermediate positions, which Nishigauchi and Fujii (2006) adduces for the evidence for movement.

(15) A: Minna$_i$-wa   [Ken$_j$-ga   dare-o   sonkeisiteiru   to] iimasita ka?
         Everyone-TOP   Ken-NOM   who-ACC   respect   C said   Q
         'Who did everyone say that Ken respects *t*?'
      B: a. Mother Teresa-o   desu ka?
             Mother Teresa-ACC   COP   Q
            'Is it Mother Teresa?'
          b. Zibun-zisin$_{i/j}$-no   hahaoya-o   desu   ka?
             Self-GEN   mother-ACC   COP   Q
            'His/her mother?'

What is interesting is that the reflexive *zibun-zisin* 'self' is either bound by the matrix subject *minna* 'everyone' or the embedded subject *Ken*. Considering that the reflexive can only be bound by the local antecedent due to

Condition A, the former reading is mysterious. This is accounted for if focus movement is successive cyclic and the reflexive is bound in the embedded CP Spec position.

(16) [[self$_{i/j}$'s mother] ~~everyone$_i$ said [$_{CP}$ [self$_{i/j}$'s mother] Ken$_j$ respects [self$_{i/j}$'s mother]]~~]]

If movement were not involved, however, it would be hard to account for binding by the matrix subject.

3 Island Insensitivity

It is known that Japanese short answers are insensitive to complex NP constraints (CNPC) except for multiple foci (Nishigauchi 1990). Regarding their island insensitivity, at least three different approaches have been proposed so far.

(17) a. Bare copular analysis (Saito 2004)
          pro/sore   XP   da/desu   or   [$_{NP}$~~[$_{TP}$... *pro* ...] no]~~ wa   XP-da/desu
      b. PF island repair analysis (Merchant 2004, Nishigauchi and Fujii 2006, a.o.)
          [$_{ForceP}$ [$_{TopP}$ [$_{FocP}$ XP$_i$ ~~[$_{FinP}$ [$_{TP}$ — ... t_i ...] no]~~ da/desu]]] (focus movement + deletion)
                                   or
          [$_{ForceP}$ ~~[$_{TopP}$ [[... t_i ...] no]$_j$~~ wa [$_{FocP}$ XP$_i$ t_j]]] (cleft + deletion)
      c. In-situ deletion analysis (Kimura 2010, Abe 2015, Kimura and Narita 2021, a.o.)
          ~~[$_{CP}$ C [$_{TP}$ — ...~~ XP da/desu ~~...]]~~

For the bare copular analysis, the island insensitivity does not cause any problem as movement is not involved, and the subject is either *pro* or its overt counterpart *sore*. In other words, it is predicted that *pro/sore* is available wherever there is a CNPC. However, it has also been noticed in the literature (Saito 2004, Nishigauchi and Fujii 2006) that *pro/sore* is not always available depending on environments, which led them to assume two sources. The in-situ deletion analysis deletes everything but a focused phrase in the base position. But as we just saw, the in-situ deletion analysis would face a difficulty in accounting for binding intermediate positions.

   I have argued that the underlining structure of the second part of Japanese split questions is a cleft sentence with the presuppositional CP subject deleted by argument deletion. So, in the same way as the cleft+deletion analysis of short answers in (17b), ellipsis salvages otherwise unacceptable island violations. I am not yet certain as to whether two sources, i.e. bare copular sentence and cleft sentence, are needed because having two of them may lead to redundancy.

   In order to illustrate this point, let us see the case of a complex NP. As shown in (18), the split question is allowed in (18Ba) but the corresponding cleft sentence is not, as in (18Bb).

(18) A: Anata-wa   [dare-ga   kaita   hon]-o   yonda no   desu ka?
         You-TOP   who-NOM   wrote   book-ACC   read   C   COP   Q
         "Lit: Did you read a book [who wrote t]?'
      B: a. (Sore-wa) Salinger desu ka?
             It-TOP   Salinger COP   Q
             'Is it Salinger?'
          b. *[Anata-ga [e_i kaita   hon]-o   yonda   no]-wa Salinger   desu ka? (Cleft)
                You-NOM   wrote book-ACC   read   C-TOP   Salinger   COP Q

'Is it Salinger$_i$ that you read the book e_i wrote?'

The PF island repair analysis would argue that the split question in (18Ba) is possible because the deletion of the presuppositional CP subject in (18Bb) repairs the island violation. On the other hand, the bare copular analysis would argue that the base structure of (18Ba) is *pro/sore XP desu ka*, which is evident in the optional use of *sore*. So, it is difficult to tell which analysis is correct just by looking at (18).

It is important to note that multiple foci, functional and pair-list questions are the environments which could separate these two analyses. Here, I only take multiple foci as an example. The following is an example adapted from Nishigauchi (1990) and changed into a split question.

(19) A: [$_{NP}$ [$_{TP}$ Dare-ga   dare-ni   kaita] tegami]-ga   mitukarimasita   ka?
           Who-NOM who-DAT   wrote letter-NOM   was.found   Q
      '[A letter [that who wrote to whom]] was found?'
    B: a. *(Sore-wa) Tanaka-san-ga   Sato-san-ni   desu   ka?
          It-TOP   Tanaka-Mr./Ms.-NOM   Sato-Mr./Ms.-DAT COP   Q
        'Is it [Mr./Ms. Tanaka to Mr./Ms. Sato]?'
      b. *[e_i e_j kaita]tegami]-ga   mitukatta] no]-wa   Tanaka-san-ga   (Cleft)
              wrote letter-NOM   was.found C-TOP   Tanaka-Mr./Ms.-NOM
          Sato-san-ni   desu   ka?
          Sato-Mr./Ms.-DAT COP   Q
        'Is it [[Mr./Ms. Tanaka$_i$ to Mr./Ms. Sato$_j$] that a letter was found [e_i e_j wrote]?'

(19Bb) shows that clefting out of a complex NP leads to an island violation and (19Ba) shows that neither PF island repair nor the replacement with *sore* 'it' salvages the island violation. So something other than these two analyses would be necessary to account for the ill-formedness of (19Ba). Note that *sore* 'it' is available even when there is no island. (20) is a sentence in which multiple foci are allowed out of a non-island CP.

(20) [$_{CP}$ [$_{TP}$ Ken-wa   dare-ni   nani-o   ageta]   no]   desu   ka? (Sore-wa) Mari-ni   ringo-o
          Ken-TOP   who-DAT   what-ACC   gave   C   COP   C   It-TOP   Mari-DAT apple-ACC
      desu   ka?
      COP   Q   'To whom and what Ken gave? Is it to Mari an apple?'

As far as these data are concerned, therefore, the PF island repair analysis is sufficient enough and the bare copular analysis may be redundant. However, much remains to be done as to whether the PF island repair analysis can cover other data.

References

Abe, J. 2015. *The In-situ Approach to Sluicing.* Amsterdam: John Benjamins. **Arregi, K. 2010.** Ellipsis in split questions. *Natual Language and Linguistic Theory* 28: 539-592. **Fiengo, R., and R. May.** 1994. *Indices and Identity.* Cambridge, MA: MIT Press. **Hiraiwa, K. and S. Ishihara. 2012**. Syntactic metamorphosis: Clefts, sluicing, and in-situ focus in Japanese. *Syntax* 15: 142-180. **Kim, S.-W. 1999**. Sloppy/strict identity, empty objects, and NP ellipsis. *Journal of East Asian Linguistics* 8: 255-284. **Kimura, H. 2010**. A *wh*-in-situ strategy for sluicing. *English Linguistics* 27: 43-59. **Kimura, H., and H. Narita. 2021**. Compound *wh*-questions and fragment answers in Japanese: Implications for the nature of ellipsis. *Linguistic Inquiry* 52: 195-209. **Kizu, M. 2005**. *Coeft Constructions in Japanese Syntax.* New York: Palgrave. **Merchant. J. 2001**. *The Syntax of Silence.* Oxford: Oxford University Press. **Merchant, J. 2004**. Fragments and ellipsis. *Linguistics and Philosophy* 27: 661-738. **Nishigauchi, T. 1990**. *Quantification in the Theory of Grammar*. Dordrecht: Kluwer Academic Publishers. **Nishigauchi, T., and T. Fujii. 2006**. Where the island remains (unrepairable). Ms. Kobe Shoin University and University of Maryland, College Park.**Oku, S. 1998.** *A Theory of Selection and Reconstruction in the Minimalist Program*. Doctoral dissertation, University of Connecticut. **Saito, M. 2004**. Ellipsis and pronominal reference in Japanese clefts. *Nanzan Linguistics* 1: 21-50.