The position of the interrogative clitic in Turkish: a Word Grammar account*

Taiki Yoshimura**

1. Introduction

In this article, I shall discuss the morpho-syntactic behaviour of the interrogative clitic *mI* in Turkish. It is well known that the interrogative clitic *mI* is positioned not only in the sentence-final, but also in a sentence-middle position. The explanation of the position of the interrogative clitic is, however, not easy, since there are cases where the sentence becomes ungrammatical because of its position. Although *mI* can appear between the genitive noun and its head-noun, the sentences become ungrammatical if *mI* is positioned immediately after the classifier noun (i.e. a noun which immediately precedes its head-noun with no genitive case-marking) as in (1b), the attributive adjective (1c), the complement-noun of the postposition (1d), and the connective clitic *dE* (1e).

(1a)   *Ahmet*   *Berna'nın **mı***   *günlüğ-ü-nü*   *bul-du?*
          A.-NOM   B.-Gen Q   diary-3SG-ACC find-PAST
          'Is it Berna's diary that Ahmet found?'

(1b)   **Ahmet*   *ders **mi***   *kitab-ı-nı*   *kaybet-ti?*
          A.-NOM   course Q   book-3SG-ACC lose-PAST
          (No reading)

(1c)   **Ahmet*   *kalın **mı***   *kitab-ı*   *oku-yor?*
          Ahmet-NOM   thick Q   book-ACC   read-PROG
          (No reading)

(1d)   *Tuğba*   *Oya-yla mı / *Oya mı-yla sinema-ya*   *git-ti?*
          Tuğba-NOM   O-COM Q / Oya Q-COM   cinema-DAT   go-PAST
          'Did Tuğba go to cinema with Oya?'

(1e)   *Ayşe de mi / *Ayşe mi de*   *cevap*   *ver-di?*
          Ayşe-NOM too Q /*Ayşe Q too   response   give-PAST
          'Did Ayşe responded too?'

The difference in grammaticality between (1a) and (1b,c) comes from the fact that the genitive noun is syntactically independent of its head-noun (i.e. possessive noun marked with possessive suffix), which has been pointed out in previous work (e.g. Hayasi 1997 and Besler 2000). The examples (1d) and (1e), on the other hand, violate some morphological rules (i.e. morphotactic rules) rather than those in syntax. In this article, I argue that the position of *mI* is controlled by both morphology and syntax, and that Besler 2000 and Aygen 2007 are attractive accounts but both have several problems to be solved, especially morphological structure involving the interrogative clitic, which is well explained within the framework of Word Grammar (WG).

2. Phrase structure analyses

2.1. Besler 2000

Besler 2000 proposes that *mI* in Turkish is a syntactic unit and therefore has syntactic properties, and suggests an analysis based on the Minimalist Program framework. According to Besler 2000, phrase structure concerned with the occurrence of *mI* (in her analysis, the question particle) is summarized as in Figure 1 below.
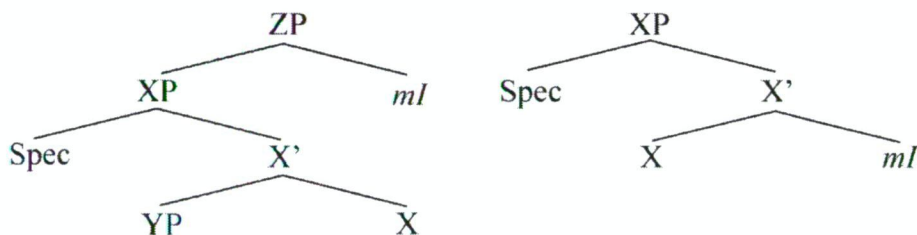

Figure 1

According to Besler 2000, when *mI* occurs with NPs, PPs, it is base-generated as a sister to these phrases, which is shown on the left side of Figure 1. On the other hand, when *mI* occurs with the verbal complex (in her terms; i.e. when *mI* occurs in the sentence-final verbal predicate, whether it precedes the agreement marker or not), in contrast to the cases of NP/PPs or (non-finite) VPs, it is not base-generated at the head of Complementizer Phrase (CP) but behaves as a suffix on heads, and it demonstrates the so-called 'LF pied-piping' with V and inflectional heads in order to turn the constituents into the scope of question, as the right side of Figure 1 above shows. In short, the question particle *mI* exhibits a 'dual behaviour', being either a lexical item or a mere suffix. The analysis is undoubtedly attractive, in that it deals with virtually all the positions where *mI* can occur in the sentence, which is advantageous over other work such as Sezer 2001 and Kornfilt 1996.

Nevertheless, I would like to point out that her analysis has at least several problems to be solved. The first issue lies in her treatment of *mI* as a linguistically separate element which varies according to its position of the occurrence or according to the constituents

to which it attaches. That is, it is treated as an independent lexical item in cases where it attaches to NPs and PPs, while it is treated as a mere suffix when it attaches to Agr or Tense/Aspect head. Consequently, it undermines the difference of the morphological properties between the clitics and the inflectional suffixes (see Erdal 2000). Secondly, as she herself notes on Besler 2000, it is still unclear whether it is base-generated as an adjunct attached to the maximal projection or as a head which has its own maximal projection (e.g. Focus Phrase) when it is base-generated as a sister to a maximal projection. That is, she leaves the problem of syntactic relation between the interrogative clitic and the word it attaches to. Finally, as she also points out by herself, her analysis does not provide the morphological device for the explanation of the relative order between the predicate with any TAM suffixes, the interrogative clitic, and the personal ending. Accordingly, it does not explain why the personal ending follows the interrogative clitic in some cases while in the other the personal ending must precede the interrogative clitic (see 5.3 below).

2.2. Aygen 2007

Aygen 2007 is another outstanding work based on the assumption that there is a 'null' (i.e. covert) Q-particle which only realizes in an echo question and a yes-no question. She assumes that this covert Q-particle is focus driven, i.e. its movement is motivated by the focus feature, and Q-particle 'covertly' moves from a sister-position to WH-word (or focused constituents) to the clause periphery (i.e. the head of CP) with the verbal complex, and surfaces only when it is necessary (in order to show the focus). This line of analysis also seems explanatory in at least two respects. Firstly, it explains the reason why *mI* does not occur in ordinary WH-questions. Secondly, it tries to generalize the characteristics of the interrogative sentences in Turkish by taking into consideration other WH-in-situ languages such as Sinhala and Japanese, where there are also Q-particles but their movements are either covert or overt.

For all her well-developed analysis, her analysis also has the following theoretical problems to be solved. First of all, as well as Besler 2000, it does not explain morphological facts about *mI*: it is the interrogative clitic which is different from both ordinary words and ordinary affixes in some respects. Second, contrary to Besler 2000, it does not explain ungrammaticality due to the position of the interrogative clitic; for example, her account, as such, does not answer the question why sentences such as (1b-d) are ungrammatical, where *mI* occurs within the NPs or PPs. It would be possible to assume that Aygen 2007 follows the ideas of Besler 2000, where *mI* is base-generated as a sister to NPs or PPs and it cannot move downward. Even so, however, then the analysis also has to answer the same kind of question that I address to Besler 2000: whether it is a syntactically independent lexical item or an inflectional suffix. In fact, Aygen 2007 assumes that the Q-particle is possible to move from a sister-position to WH-word (rather than a sister to NPs or PPs) to the clause-peripheral, so Aygen 2007 cannot assume the interrogative clitic in sentence-middle to be a lexical item on the one hand, and it becomes an inflectional suffix when it surfaces with the verbal complex on the other hand. Otherwise, it is necessary to presume that *mI* changes its formal status, i.e. from a lexical item to a suffix along the way to the clause-periphery position, which does not seem to be well-motivated.

In summary, both Besler 2000 and Aygen 2007 seem to be explanatory works which aim to predict all the possible position of the interrogative clitic in an integrated way, but unfortunately both have several problems including morphological aspects of the interrogative clitic.

3.   A Word Grammar account

WG is a theory for language structure which has been proposed by Hudson (2007, 2010, among others). The main characteristics of WG for our purpose here are summarized as follows. First, it presents language as a network of knowledge, linking concepts about words, their meanings, etc. For example, a word DOG is linked to the meaning 'dog', to the form /dAg/, to the word-class 'noun', etc. Second, WG assumes every sentence structure to be monostratal. That is, only one structure per sentence is assumed, hence no transformations (or movement) is necessary for syntax. Third, WG uses word-word dependencies rather than phrase structure; for example, a noun depends on a verb as its subject. And finally, it shows grammatical relations/functions by explicit labels such as 'subject' and 'object'. In what follows, we shall see some more details of this framework for the purpose of our discussion.

The first notion introduced for the discussion is the notion 'landmark' (hereafter lm). This is a temporal and/or spatial relation between concepts (including words) which any person has. This is a general concept, but WG assumes that this is also applicable to language structure, in that a word takes its position from another word. WG also assumes that in principle, a word's landmark is its syntactic head (in WG's term, 'parent'). For example, in *good books* in English, *books* is the parent of its dependent *good*, so *books* is the landmark of *good*. Similarly, taking *kalın kitap* in Turkish for example, *kitap* is the landmark of *kalın*.

A problem about word order arises when we consider examples like **Good read books.* in English, and **Kalın okuyor kitabı.* in Turkish, where all words have proper landmarks but as a result it may make discontinuous word order. These cases indicate that just to prescribe individual landmark relation between two words is not enough. For this, WG assumes that landmark relation is transitive, which is formulated as (2):

(2)   The order concord (also known as 'Landmark Transitivity' (Hudson 2007: 139)):
         If X is the landmark of Y and Y is before X, and Y is the landmark of Z, then Z
         is also before X (and similarly for 'after')

The principle in (2) applies to the ungrammatical examples shown above, which is shown in the left side of Figure 2. The notation is, however, more or less complicated for our current purpose; so, an easier notation for explaining continuity of word order is introduced in the current WG framework. That is, dependency arrows must not tangle (Hudson 2010), along with recognizing 'sentence-root' arrows which point to the word which has no syntactic parent in the sentence. This easier notation is illustrated the right side of Figure 2.
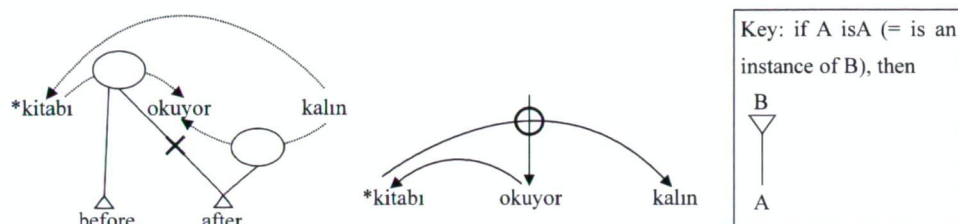
Figure 2

Note that the No-tangling Principle only applies to dependency arrows at surface (i.e. arrows drawn above each word in the right side of Figure 2) which also involve landmark relation. In any case, WG blocks inappropriate word order for discontinuity by a rule which requires syntactically related words to be adjacent to each other.

Another principle that is necessary to introduce the No-tangling Principle rigidly is the Raising Principle (hereafter RP). This principle comes from cases where a word has two syntactic parents, as in *John is coming* in English or *Ben yardımcı olurum* (I-Nom helper be-Aor-1sg; literally 'I will help you') in Turkish. The problem here is which words are the landmarks in these examples. So, WG provides another principle called the Raising Principle, which is shown in (3) below:

(3)   The Raising Principle: if a word has more than one parent, then its landmark is the parent which is superordinate to all the other parents. (Hudson 2007: 141)

Therefore, we illustrate the structures of our stored examples *Ben yardımcı olurum* as shown at the left side, rather than one at the right side of Figure 3:
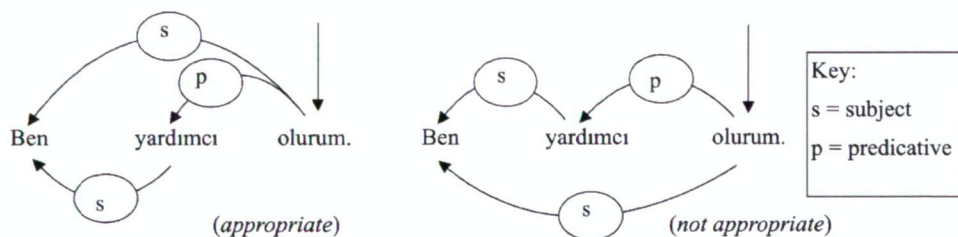

Figure 3

In summary, the following notions concerned with word order have been introduced: the landmark, the landmark transitivity and the Raising Principle. Within these generalizations about word order of WG, we shall try to explain syntactic behaviour of the genitive noun and the interrogative clitic *mI* in the next section.

4. The analysis

4.1. Comparison: genitive noun phrase and other noun phrase

As we have seen in Section 1, we assume that the genitive noun is relatively independent of its head noun in terms of word order, whereas other noun modifiers are not. In terms of WG, the landmark of the genitive noun is not its head-noun, but the predicate of the sentence. If this is so, the genitive noun has two landmarks (i.e. syntactic parents), namely the head-noun marked with the possessive suffix and the predicate itself. It is helpful to recall that only the genitive noun can be extraposed into the post-verbal position, which means that the location of the genitive noun is controlled by the predicate of the sentence.

If this is correct, then the answer to the question about the position of *mI* is obvious: the interrogative clitic depends on the predicate of the sentence syntactically, so the predicate is by definition the landmark for *mI*. Thus, the genitive noun and *mI* are co-dependents of the predicate of the sentence. Furthermore, if Turkish speakers know (even unconsciously) that the genitive noun is the 'after' landmark for the interrogative clitic (and vice versa), then the example (1a) *Ahmet Berna'nın mı günlüğünü buldu?* is predicted to be grammatical because it does not violate any syntactic rule, especially No-tangling Principle. By contrast, examples (1b) and (1c), which contain a modifier, not a genitive noun, attached to the head-noun are all predicted to be ungrammatical, because they all violate the No-tangling Principle. All the discussion so far is illustrated in Figure 4 below:
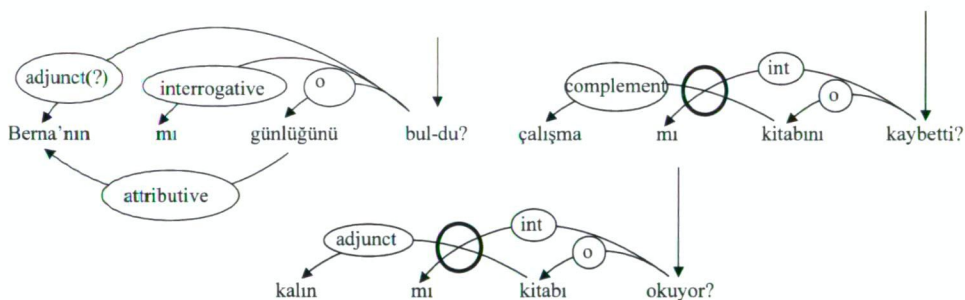

Figure 4

As Figure 4 shows, only the genitive noun depends on the possessed noun. So, it allows the interrogative clitic *mI* to immediately follow it. In addition, the analysis proposed in this subsection dispenses with any abstract syntactic operation like covert movement or 'LF pied-piping'; hence it can also handle the problem concerned with the interrogative clitic without mixing into the problem of scrambling. In this sense, our explanation has an advantage over previous analyses by Besler 2000, where *mI* is assumed to be an independent word in sentence-middle on the one hand, but as an inflectional suffix in sentence-final on the other hand, which should otherwise be justified in some ways.

4.2. Morphological explanation

One of remaining problems is how to deal with the clitic cluster in a sentence-middle position. Let us return to ungrammatical examples shown in (1d) and (1e), where *mI* and another clitic such as *ylA*, the comitative clitic, and *DA*, the connective clitic, are used in combination, and the sentence becomes ungrammatical unless the clitics are combined in a certain order. What is crucial here is that, in both cases, *mI* is in a sentence-middle position, and resorting to some syntactic explanation alone is not sufficient to account for them. Therefore, some other theoretical framework than the syntactic one is needed. WG provides a rich network for morphological structure whose representation is quite similar to its syntactic dependency network.

In WG, words and their forms are clearly distinguished, and there are at least two types of form which are provided for the explanation of clitics: the 'clitic-form' and 'host-form' (Hudson 2007). The clitic-form is, by definition, the form of a clitic, and the host-form is a special kind of word-form which contains at least one clitic, along with the clear distinction between word-level and form-level. The analysis is shown in Figure 5 below, where forms are represented by curly brackets:
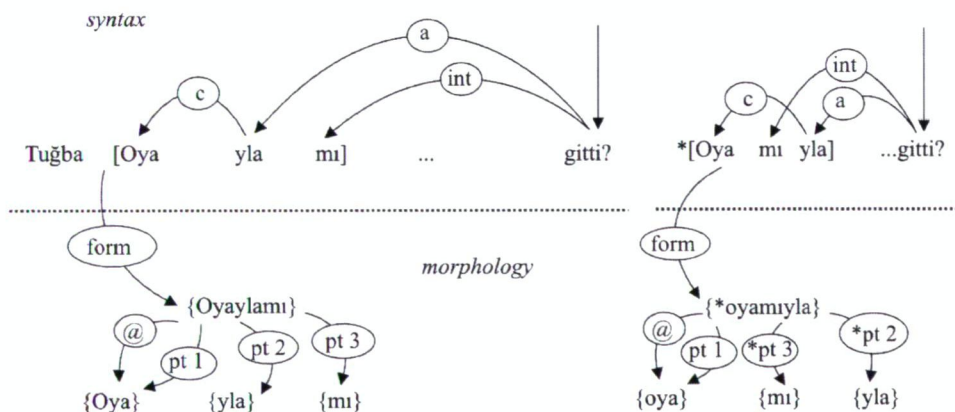

Figure 5

In Figure 5 above, symbols such as 'a', 'c', and 'int' written in bubbles with curved arrows indicate 'adjunct', 'complement' and 'interrogative' relations between words concerned respectively. Accordingly, *Oya*, *yla*, and the interrogative clitic are recognized as words in syntax; by definition, each word is syntactically independent of other words, and syntactic dependencies (e.g. 'adjunct' and 'complement') between these words should be drawn above these words.

By contrast, in morphology *{Oyaylamı}* is the host-form (which is in turn a special kind of word-form) which contains at least one clitic-form (i.e. *{mı}* shown in Figure 5 above). In this form-level, the label '@' is the abbreviation for the 'anchor' (Hudson 2007), showing that the host-form *{Oyaylamı}* takes its position from *{Oya}*, and those labelled as 'pt' with number in bubbles also show 'part' relations in the host-form. So, Figure 5 above shows that *{Oya}*,*{yla}* and *{mı}* are 'part1', 'part2' and 'part3' of the host-form *{Oyaylamı}*.

This numbering is important because these relations are responsible for the rigid order of suffixes and clitics within a word, which we shall discuss this morphotactic rules in the next subsection.

In summary, the ungrammatical forms as in (1d) (1e) violate (at least) morphotactic rules within the host-forms. We see that morphology plays an important part in explaining the position of the interrogative clitic, as well as syntactic dependencies between words concerned. Morphology is also important because, however indirectly, we can explain why the interrogative clitic undergoes the so-called Vowel Harmony Effect, regardless of whether it appears in sentence-middle or sentence-final.

4.3. The interrogative clitic in sentence-final

I argue that our analysis proposed also applies to cases where *ml* is positioned at sentence-final (Sezer 2001, Kornfilt 1996). It is a well known fact that almost all the affixes follow their stem and the order of these affixes is rigidly fixed. As examples (4a,b) show, in the relatively simple example *gel-di-m* 'I came', all two suffixes {ti} (the past tense suffix) and {m} (the first singular suffix) must follow its stem {gel}, and the order of these forms are rigidly fixed, which is thought to be maintained by some kind of morphotactic rule.

(4a)   a. stem (pt1)   TAM: definite past (pt2)   person (pt3)
        gel-   *di-*   *m*

(4b)   stem (pt1)   person (*pt3)   TAM: definite past (*pt2)
        *gel-   *m-*   *di*

The same is true of the cases where inflectional suffixes as well as clitics are concerned. In (5a), a host-form contains the personal inflectional suffix. It is ungrammatical when the interrogative clitic precedes a personal suffix (i.e. **okudu mun?*) or when the paradigm of personal ending is not correct such as **okudu munuz?* Similarly, we can predict the appropriate host-form in the cases such as (5b), where ungrammatical forms show either the wrong order of the two clitics (i.e. **okuyorsun mu?*) or disagreement between the type of the personal form and the Tense/Aspect/Mood form (i.e. **okuyor mun?*}.

(5a)   *Sen*   *kitab-ı*   *oku-du-nuz mu? /*oku-du mu-n? /*oku-du mu-sunuz?*
        you-Nom book-Acc   read-Past-2pl Q read-Past Q-2sg read-Past Q-2pl
        'Did you read the book?'

(5b)   *Sen*   *kitab-ı*   *oku-yor mu-sun? /*oku-yor-sun mu? /*oku-yor mu-n?*
        you-Nom book-Acc   read-Prog Int-2sg read-Prog-2sg Int read-Prog Int-2sg
        'Are you reading the book?'

The WG analysis for cases where the relative order between the TAM suffix, the interrogative clitic, and the personal endings are relevant is shown in Figure 6:
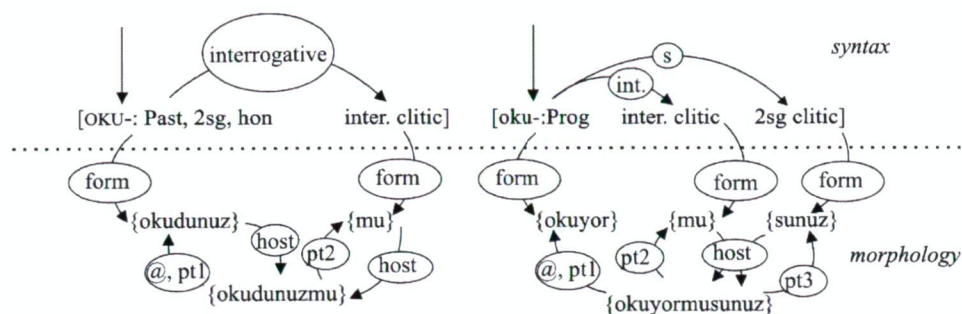
Figure 6

As Figure 6 shows, the 'host-form' analysis is also available for cases where the interrogative clitic appears at the sentence-final position. The analysis is quite similar to those in Figure 5.

I assume here that the interrogative clitic is not the head of the predicate for several reasons (Yoshimura 2011).[1] If *mI* is recognized as the dependent of the predicate, then we can say that the position of the interrogative clitic is syntactically always headed by the predicate, regardless of whether it appears in sentence-middle or sentence-final (for more detail of the discussion, see Yoshimura 2011).

It is fair to say that our analysis has advantages over previous analyses shown in Section 2, namely that it is the morphotactic rule that determines the position of the interrogative clitic in the cases where *mI* occurs with the verbal complex: if there is a pronominal clitic in Turkish, then it is required to follow the interrogative clitic by some morphotactic rule, and if it is a personal (i.e. inflectional) suffix it has to precede the interrogative clitic.

5.   Conclusion

So far, I have shown that a WG account explains the position of the interrogative clitic without assuming any movement rule as Besler (2000) and Aygen (2007) assume. A possible objection to my analysis is that a WG account does not explain the scope of question. For lack of space, I will not go into the discussion about semantic structure in WG which is concerned with the scope of question. All I can say at present, however, is that in fact it is implicit both in morphology and syntax, although it is pointed out that there are exceptional cases where the position of *mI* and the scope of question do not correspond to each other (e.g. Zimmer 1997, Göksel and Kerslake 2005). That is to say, it is the host-form which contains the interrogative clitic-form, and simultaneously it is the word positioned immediately before the interrogative clitic that is placed into the scope of question.

---

1 For more detail, see Yoshimura 2011, which regrettably contains misunderstanding of the ideas of Besler 2000.

References

Aygen, G. 2007. Q-Particle. *Journal of Linguistics and Literature* 4: 1, 1–30.

Besler, D. 2000. *The question particle and movement in Turkish*. Unpublished MA thesis, Boğaziçi University.

Erdal, M. 2000. Clitics in Turkish. In: Göksel, A. & Kerslake, C. (eds.) *Studies on Turkish and Turkic languages*. Wiesbaden: Harrassowitz. 41–48.

Göksel, A. & Kerslake, C. 2005. *Turkish: a comprehensive grammar*. London: Routledge.

Hayasi, T. 1997. Separated genitive construction in Modern Turkish. In: Matsumura, K. & Hayasi, T. (eds.) *The Dative and Related Phenomena*. Tokyo: Hituzi Shobo. 227–253.

Hudson, R. 2007. *Language networks: the New Word Grammar*. Oxford: Oxford University Press.

Hudson, R. 2010. *An introduction to Word Grammar*. Cambridge: Cambridge University Press.

Kornfilt, J. 1996. On copular clitic forms in Turkish. *ZAS Papers in Linguistics* 6. Berlin: ZAS. 96–114.

Yoshimura, T. 2009. Does the interrogative clitic *mi* in Uzbek head the predicate? In: Kang Yong-Se et al. (eds.) *Current Issues in Linguistic Interfaces* 1. Seoul: Hangkookmunhwasa. 235–246.

Yoshimura, T. 2011. Syntactic Independence of the Genitive Noun and the Position of the Interrogative Clitic in Turkish: A Word Grammar Account. In: Sugayama, K. (ed.) *Kyoto Working Papers in English and General Linguistics*. Tokyo: Kaitakusha. 3–28.

Yoshimura, T. (in preparation) *A Word Grammar Account of the grammatical behaviour of the interrogative clitic in Turkish*.

Zimmer, K. 1997. The case of the errant question marker. In: Johanson, L. et al. (eds.) *The Mainz Meeting*. Wiesbaden: Harrassowitz. 478–481.