Mermaid Constructions in Lexical Functional Grammar

Chihiro Taguchi
University of Notre Dame

Proceedings of the LFG'22 Conference

Miriam Butt, Jamie Y. Findlay and Ida Toivonen (Editors)

2022

CSLI Publications

http://csli-publications.stanford.edu/

Abstract

This paper provides a cross-linguistic analysis of Mermaid Constructions in terms of Lexical Functional Grammar. Mermaid Constructions, coined by Tsunoda (2020), are grammaticalized monoclausal constructions in which a verb and a noun, sometimes with a copula, form a compound predicate. However, the work was chiefly descriptive, and the morphosyntactic nature of Mermaid Constructions in theoretical terms has not yet been explained. In this work, in opposition to Tsunoda's (2020) hypothesis, the following points are shown: (1) Mermaid Constructions are not monoclausal but biclausal; (2) Mermaid Constructions do not comprise a compound predicate, but are control and raising with a nominal predicate; (3) these findings hold cross-linguistically.

1 Introduction

The term "Mermaid Construction" (MC henceforth) was coined by the linguist Tasaku Tsunoda (2011, 2012, 2020). MCs refer to constructions in which a verb and a noun (and also a copula, depending on the language) exhibit a close-knit morphosyntactic relationship, which, according to Tsunoda (2020), form a compound predicate. The nouns that can trigger this construction (MC nouns henceforth) are limited. As for their meaning, the semantic properties of the noun have been somewhat grammaticalized into a functional meaning such as attitude, aspect, modality, and evidentiality. Sentences (1) and (2) are canonical examples of MCs with the MC nouns *ki* ("feeling") and *yotei* ("plan") from Japanese.[1] Among the languages that have MCs, Japanese is reported to have the highest number of MC nouns (Tsunoda 2020, p. 12). Although the literal translation of (1) would be composed as either "Hanako is the feeling that goes to the UK" (assuming that *Hanako* is a subject of the matrix clause) or "It is the feeling where Hanako goes to the UK" (assuming that *Hanako* is the subject of the embedded relative clause), its correct free translation is "Hanako intends to go to the UK," as if it had an attitude predicate. Similarly, (2)'s free translation differs from its possible literal translations as if *yotei=da* functioned as a prospective aspectual marker.

[†]I deeply appreciate constructive and significant feedback from the students, the linguists, and the anonymous reviewers to my work which was originally presented at the 27th International Lexical Functional Grammar Conference (LFG22) and the 33rd South of England LFG Meeting (SELFG33). I would also like to thank native speakers who validated the sentences used in the paper: Röstäm Sülti, Bulat Şäymi, Kamilä Raqıypova, and Aysulu Saetgaraeva for Tatar, Rae Pineda and Lean Yao for Tagalog, and Yasmein Khalil for Egyptian Arabic. This material is based upon work supported by the National Science Foundation under Grant No. BCS-2109709. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation.

[1]The list of abbreviations used in glossing is attached to the end of this paper.

(1) *Hanako=ga Igirisu=ni ik-u ki=da*
Hanako=NOM UK=DAT go-NPST.ADN feeling=COP.NPST.DECL
'Hanako intends to go to the UK. (*lit.* "Hanako is the feeling to go to the UK" or "It is the feeling where Hanako goes to the UK")'

(2) *Hanako=ga Igirisu=ni ik-u yotei=da*
Hanako=NOM UK=DAT go-NPST.ADN plan=COP.NPST.DECL
'Hanako is going to go to the UK. (*lit.* "Hanako is the plan that goes to the UK" or "It is the plan where Hanako goes to the UK")'

A striking characteristic of MCs is that (i) an MC noun is a morphosyntactically fully realized noun; however, (ii) an MC noun seems to partially play the role of a non-canonical nominal predicate, i.e., the nominal predicate does not function as a typical copular nominal predicate such as an equative (e.g., "The Morning Star **is the Evening Star**") or a subset relationship (e.g., "Cats **are animals**"); in other words, (iii) an MC noun seems to semantically contribute to rather abstract meanings such as aspect, modality, and evidentiality. In this respect, MCs are an intriguing grammatical phenomenon that exhibits morphosyntactically, functionally, and semantically idiosyncratic usage.

As the title of Tsunoda (2020) succinctly states, previous work claims that MCs have a monoclausal structure whose predicate is a compound predicate. Therefore, the verbal predicate *ik-u* and the nominal predicate *yotei=da* in (2) form one constituent as a compound predicate, as shown with a double underline in (3).

(3) Japanese (< Japonic; SOV)
Hanako=ga Igirisu=ni <u>ik-u+yotei=da</u>
Hanako=NOM UK=DAT go-NPST.ADN+plan=COP.NPST.DECL
'Hanako is going to go to the UK.'

As the prototypical characteristics of MCs, Tsunoda (2020, p. 4) tentatively lists five common properties shown in (4). (4a) states that the structure superficially looks as if the Clause modifies the nominal predicate by some relation, as can be observed in (2). Note that a copula can be absent in some languages and that the actual word order is not restricted to that of (4a). (4b), though definitions of a word and a clitic are not given, characterizes that the Noun is at least homophonic with a noun used elsewhere, that is, not phonologically reduced. For instance, the Japanese MC noun *yotei* is homophonic with the noun that means a "plan". (4c) stands on the assumption that the subject belongs to the (seemingly embedded) Clause, not the matrix clause, which will be rejected in the following arguments; for now, if we accept this view for (2), it is clear that they are not co-referential, because Hanako is an individual and is not a plan. (4d) states that an MC's Clause can be a grammatical sentence on its own without the MC noun and copula. In (1), for example, *Hanako=ga Igirisu=ni ik-u* without the Noun and Copula *ki=da* is also a grammatical sentence. However, this is a controversial generalization and

will be rejected in later discussions. The fifth generalization (4e) excludes constructions where the Clause is either nominalized or complementized and stands as the subject of the nominal predicate. For example, in Japanese (2), "That Hanako goes to the UK is the plan" is not a possible syntactic interpretation.

(4) (a) The superficial structure is [Clause] Noun Copula.
(b) The Noun is an independent word, not a clitic.
(c) The subject of the Clause and the Noun are not co-referential.
(d) The Clause can be used as a sentence by itself.
(e) The Clause is not the subject of the Noun + Copula.

The following sections in this paper will propose an alternative analysis of MCs. In particular, rejecting the assumptions given by Tsunoda (2020), the following points will be argued for. First, from the descriptive evidence, we demonstrate that assuming a compound predicate is not an essential syntactic device to account for MCs. Specifically, it will be shown that several languages, which were not mentioned in Tsunoda (2020), turn out to be counterexamples that express MCs without linear adjacency of a verbal predicate and an MC noun. Second, to provide an alternative explanation, we propose a biclausal analysis of MCs as opposed to the monoclausal hypothesis claimed by Tsunoda (2020). Third, we further argue that MCs involve control and raising with a nominal predicate. For these two claims, data from Japanese are used as the primary evidence to keep the argument in line with Tsunoda's (2020). Finally, we extend this analysis to other languages with MCs for the sake of cross-linguistic generalization. For the detailed syntactic analysis, we employ the Lexical Functional Grammar (LFG) framework, because its architecture that severs a grammar into separate modules such as a c(onstituent)-structure and an f(unctional)-structure successfully captures the MC noun's peculiar status being a morphologically independent noun yet semantically contributing to aspect, modality, evidentiality, and attitude.

Note that this paper only considers MCs where the nominal component is a full noun, and not where it is a clitic, affix, or adjectivalized noun, as in the 'periphrastic' varieties of MC mentioned by Tsunoda (2020). While these latter types are clearly related kinds of MC, they diverge so greatly from canonical MCs that we believe they require a different kind of analysis.

2 Objections to the monoclausal hypothesis of MCs

This section provides counterarguments against Tsunoda's (2020) hypothesis that MCs have a monoclausal structure with a compound predicate. First, we present evidence from Welsh and Russian to succinctly demonstrate that a verbal predicate and an MC noun do not compose a compound predicate. In these languages, a verbal predicate and an MC are not always pronounced next to each other, and other syntactic constituents may intervene between them. Then, instead of assuming a

compound predicate, we propose that MCs are biclausal. In addition, we show with typological evidence that MCs most typically contain a non-finite clause that cannot stand alone as a sentence, arguing against (4d). Last but not least, at the end of this section, we point out that the observations that MCs are "uncommon crosslinguistically" and are "in the main confined to Asia" (Tsunoda 2020, p. 53) are hastily drawn generalizations by giving more examples that are not included in the work.

2.1 MCs do not form a compound predicate

Although detailed discussions as to what kind of morphosyntactic structure the compound predicate has are not provided in Tsunoda (2020), they seem to assume a linear adjacency of a verbal predicate and an MC noun as the prototypical structure of MC (cf. Tsunoda 2020, p. 35). However, MCs in languages like Welsh, Scottish Gaelic, and Russian can be expressed without their verb and noun being next to each other, as emphasized with wavy underlines in (5)–(7).[2,3] Yet, at this point, it remains possible that they form a complex predicate with a single argument structure (Butt 2014, Zaenen and Dalrymple 1995).

(5) Welsh
Rhaid i fi godi'n gynnar
rhaid i fi godi yn gynnar
necessity to 1SG wake_up.VN in early
'I need to wake up early.'

(6) Scottish Gaelic (W. Lamb, p.c.)
b' àbhaist dhi snàmh
COP.PST custom to:her swim.VN
'She used to swim.'

(7) Russian
pora nam bylo uxodit'
time 1PL.DAT COP:PST:N PFV:leave:INF
'It was time for us to leave.'

[2]One might argue that the Russian MC is merely an impersonal sentence with a nominal predicate modified by the infinitive verb, just like "It was time for us to leave." in English. However, the crucial difference is that Russian *pora* is lexically selective, i.e., does not allow substitution with a synonym or modification by an adjective, while in English it is possible to say "It was high time for us to leave." with adjectival modification.

[3]VN in the glossing for Welsh and Scottish Gaelic is a so-called "verbal noun" in Celtic linguistics (Ramchand 1993, Li 2004); however, it does not mean that its syntactic function is entirely nominalized. See the cited works for more discussion.

2.2 MCs are biclausal: Evidence from Japanese

Though the evidence shown in (5) and (7) already suggests the biclausality of MCs, this subsection presents further data from Japanese supporting a biclausal analysis, some of which are, by contrast, used as supporting evidence for the monoclausal analysis in Tsunoda (2020).

2.2.1 Negation

Partial evidence for the biclausality of MCs is that the scope of the negation is uniquely determined. While, for example, monoclausal complex predicates in Urdu have ambiguous interpretations as to the scope of the negative (Butt 1995), Japanese MCs have a clear difference as shown in (8) and (9). It is also possible to negate both the event of planning and the event of going as in (10).

(8) *Hanako=ga* [*Nagoya=ni* *ik-anai*] *yotei=da*
Hanako=NOM Nagoya=DAT go-NEG.NPST.ADN plan=COP.NPST.DECL
'It is planned that Hanako will not go to Nagoya.'
yotei(¬go(Hanako, Nagoya))

(9) *Hanako=ga* [*Nagoya=ni* *ik-u*] *yotei=dewanai*
Hanako=NOM Nagoya=DAT go-ADN.NPST plan=COP:NEG.NPST.DECL
'It is not planned that Hanako will go to Nagoya.'
¬yotei(go(Hanako, Nagoya))

(10) *Hanako=ga* [*Nagoya=ni* *ik-anai*]
Hanako=NOM Nagoya=DAT go-NEG.ADN.NPST
yotei=dewanai
plan=COP:NEG.NPST.DECL
'It is not planned that Hanako will not go to Nagoya.'
¬yotei(¬go(Hanako, Nagoya))

2.2.2 Double temporal adverbials

A monoclausal sentence cannot have two conflicting temporal adverbials. However, Japanese MCs can also contain two temporal adverbials that contradict each other as in (11). This contrasts with a monoclausal sentence with =*soo*, a reported evidential, which is possibly grammaticalized from the Early Middle Japanese MC noun *sama* (Tsunoda 2020, p. 41). Together with the argument in Section 2.2.1, this evidence shows that these MCs include two different events.

(11) ***kinoo*** *Hanako=ga* [***asita*** *Nagoya=ni* *ik-u*]
yesterday Hanako=NOM tomorrow Nagoya=DAT go-NPST.ADN
yotei=datta
plan=COP.PST.DECL
'Yesterday, it was planned that Hanako would go to Nagoya tomorrow.'

(12) ***kinoo*** *Hanako=ga* ***asita*** *Nagoya=ni*
yesterday Hanako=NOM tomorrow Nagoya=DAT
ik-u=soo=datta
go-NPST.ADN=HRS=COP.PST.DECL
'(intended:) Yesterday, (I heard) Hanako would go to Nagoya tomorrow.'

2.2.3 Honorific agreement

Although Japanese morphology lacks typical grammatical agreement like tense, person, and number, a subject that is grammatically honored by the speaker exhibits optional agreement with its predicate as shown in (13) (Matsumoto 1996, p. 27). In Japanese MCs, it is possible to use an honorific either in the MC noun (14), in the verbal predicate (15), or both (16), while honorific agreement with the predicate noun in the typical adnominal modification structure is unacceptable (17).[4] These data show that *sensei=ga* "teacher" is the subject of both *yak-u* "grill' and *yotei* "plan". If we take the monoclausal approach with a compound predicate, it remains enigmatic as to why both the verbal predicate and the MC nominal predicate can separately agree with the subject in honorification. On the other hand, the variation can be explained if we assume that (14)–(16) comprise two clauses whose subject is shared. The details as to the position of the subject and how the subject is shared will be discussed in the following parts of this section and in Section 3.

(13) *sensei=ga* *sakana=o* *oyakininar-u*
teacher=NOM fish=ACC grill:HON-NPST.DECL
'The teacher grills the fish.'

(14) *sensei=ga* *sakana=o* *yak-u* *go-yotei=da*
teacher=NOM fish=ACC grill-NPST.ADN HON-plan=COP.NPST.DECL
'The teacher is going to grill the fish.'

(15) *sensei=ga* *sakana=o* *oyakininar-u* *yotei=da*
teacher=NOM fish=ACC grill:HON-NPST.ADN plan=COP.NPST.DECL
'The teacher is going to grill the fish.'

(16) *sensei=ga* *sakana=o* *oyakininar-u* *go-yotei=da*
teacher=NOM fish=ACC grill:HON-NPST.ADN HON-plan=COP.NPST.DECL
'The teacher is going to grill the fish.'

(17) #*kore=ga* *sensei=ga* *sakana=o* *yak-u*
this=NOM teacher=NOM fish=ACC grill-NPST.ADN
o-nioi=da
HON-smell=COP.NPST.DECL
'This is the smell that [comes from where] Hanako grills the fish. ("this", or potentially, the possessor of the smell is honored)'

[4] Japanese relativization can be "gapless" (Matsumoto 1988), that is, it is possible to modify a noun with a relative clause whose verb does not subcategorize for the noun as a core argument.

2.2.4 The MC's subject is in the main clause

So far, it has been shown that the Japanese MCs are biclausal and their matrix and subordinate subjects are shared. This section argues that the pronounced subject is in the matrix clause because the subject in Japanese MCs does not allow nominative–genitive alternation that is allowed in adnominal modification constructions.

In a relative clause formed by adnominalization, it is possible to use a genitive case marked on the embedded subject as exemplified in (18). In contrast, an MC (19) does not allow the genitive case-marking but only allows nominative (19). This is used as part of the evidence for MCs' monoclausality by Tsunoda (2020). However, this fact merely shows that the subject *Hanako* in (19) belongs to the matrix clause and does not directly justify their monoclausality. The biclausal analysis can also account for the restriction on the nominaitve–genitive alternation by assuming that the pronounced subject belongs to the matrix clause, and the embedded subject is controlled by it either anaphorically or functionally. The details of the control will be discussed in Section 3.

(18) Non-MC
pro [*Hanako*={*ga/no*} *yak-u*] *sakana=da*
pro Hanako=NOM/GEN grill-NPST.ADN fish=COP.NPST.DECL
'(It is) the fish that Hanako grills.'

(19) MC
Hanako=*{ga/*no}* [*sakana=o yak-u*] *yotei=da*
Hanako=NOM/GEN fish=ACC grill-NPST.ADN plan=COP.NPST.DECL
'Hanako is going to grill the fish.'

This subsection has shown evidence for the biclausal analysis and that data adduced to argue for a monoclausal account are equally compatible with a biclausal one. The next subsection will argue against the typological tendency and the generalization (4d) mentioned by Tsunoda (2020).

2.3 Other remarks on the generalization of MCs by Tsunoda (2020)

The generalization (4d) states that "the Clause can be used as a sentence by itself" (i.e., finite). It is true that the verbal Clause in Modern Japanese can be used as a sentence by itself. However, this is due to the syncretism of the adnominal form and the declarative (finite) form in the verbal paradigm of Modern Japanese. When the predicate of the Clause is an adjectival noun (also known as *na*-adjective or nominal adjective) that retains the distinction between the adnominal and declarative forms, only the sentence with the adnominal form is grammatical (20). Other languages that have MCs with a seemingly finite Clause, such as Irabu Ryukyuan (Shimoji 2020) and Sidaama (Kawachi 2020), also have a shared form in adnominal and finite verb forms. In fact, most languages other than these express an MC with

a non-finite form, so it is more natural to claim that the verbal predicate in MCs is typically non-finite. Note that a few languages express MCs with a truly finite predicate, such as Tagalog (21) and Egyptian Arabic (22).

(20) *aki=wa* *kooyoo=ga* *kirei=*{*na/*da*}
autumn=TOP red_leaves=NOM beautiful=COP.ADN/COP.DECL
yotei=da
plan=COP.NPST.DECL
'Regarding the autumn, the red-colored foliage is going to be beautiful.'

(21) Tagalog (Katagiri 2020)
mukha[*=ng* *bi-bisita* *si=Noy* *sa=Davao*]
face=LK AF:CONT-visit NOM=Noy OBL=Davao
'It seems that Noy will go to Davao.'

(22) Egyptian Arabic (ElSadek and Sadler 2015)
Mary *šakl-aha* [*ḍarabet* *John*]
Mary shape-POSS.3SG.F hit.PFV.3SG.F John
'Mary seems to have hit John (*lit.* Mary, her shape, she hit John).'

In addition, Tsunoda (2020) points out the limited geographical distribution of MCs: "[w]ith one exception — Sidaama of Ethiopia, Africa — it has been reported only from languages of Asia, clustering in northeast Asia and in Tibeto–Burman languages, in particular" (p. vii). However, the languages with MCs shown in this paper but not shown in Tsunoda (2020), Tatar, Russian, Welsh, Scottish Gaelic, and Egyptian Arabic, are all non-Asian languages. While it is true that MCs are more frequently found in languages of Asia than those of other regions, the data provided in this paper suggest that this geographical tendency is not an areal feature and that there may be more languages outside Asia whose MCs have yet to be pointed out.

3 MCs as control and raising with a nominal predicate

The previous section provided supporting arguments for the biclausality of MCs. This section shows further evidence that MC nominal predicates trigger anaphoric and functional control employing the same diagnostics used for control and raising verbs.

3.1 Passivization test

The passivization test checks if passivizing the transitive infinitive verb and promoting/demoting their arguments accordingly changes the truth conditions (Polinsky 2013). If it does, the main predicate is a control predicate. If it does not, the main predicate is a raising predicate. This contrast comes from the difference in the argument structures of control and raising predicates, where the former directly

subcategorizes for a subject and the latter does not. For example, in English, passivization with the functional control predicate *seem* does not change the thematic relation (23), while the passivized counterpart with the anaphoric control predicate *try* has a different structure in its semantics (24). This difference arises from the argument structure of an anaphoric control predicate.

(23) (a) Tom seems to hit Jerry. $\mathsf{seem}(\mathsf{hit}(\mathsf{Tom}, \mathsf{Jerry}))$

(b) Jerry seems to be hit by Tom. $\mathsf{seem}(\mathsf{hit}(\mathsf{Tom}, \mathsf{Jerry}))$

(24) (a) Tom tries to hit Jerry. $\mathsf{try}(\mathsf{Tom}, \mathsf{hit}(\mathsf{Tom}, \mathsf{Jerry}))$

(b) Jerry tries to be hit by Tom. $\mathsf{try}(\mathsf{Jerry}, \mathsf{hit}(\mathsf{Tom}, \mathsf{Jerry}))$

Similarly, the two Japanese MC nouns *yotei* "plan" and *ki* "feeling" exhibit the same contrast. While passivization with *yotei* retains the thematic relation of the arguments as in (25), passivization with *ki* causes the "intender" to change (26).

(25) (a) Functional control, active
Hanako=ga [*Taroo=o* *tatak-u*] *yotei=da*
Hanako=NOM Taro=ACC hit-NPST.ADN plan=COP.NPST.DECL
'Hanako is going to hit Taro.' $\mathsf{planned}(\mathsf{hit}(\mathsf{Hanako}, \mathsf{Taro}))$

(b) Functional control, passive
Taroo=ga [*Hanako=ni* *tatak-are-ru*] *yotei=da*
Taro=NOM Hanako=DAT hit-PASS-NPST.ADN plan=COP.NPST.DECL
'Taro is going to be hit by Hanako.' $\mathsf{planned}(\mathsf{hit}(\mathsf{Hanako}, \mathsf{Taro}))$

(26) (a) Anaphoric control, active
Hanako=ga [*taroo=o* *tatak-u*] *ki=da*
Hanako=NOM Taro=ACC hit-NPST.ADN feeling=COP.NPST.DECL
'Hanako intends to hit Taro.' $\mathsf{intend}(\mathsf{Hanako}, \mathsf{hit}(\mathsf{Hanako}, \mathsf{Taro}))$

(b) Anaphoric control, passive
Taroo=ga [*Hanako=ni* *tatak-are-ru*]
Taro=NOM Hanako=DAT hit-PASS-NPST.ADN
ki=da
feeling=COP.NPST.DECL
'Taro intends to be hit by Hanako.' $\mathsf{intend}(\mathsf{Taro}, \mathsf{hit}(\mathsf{Hanako}, \mathsf{Taro}))$

3.2 Idiom chunk test

The idiom chunk test checks if using the given predicate in an idiomatic expression cancels the idiomatic meaning or not. If it forces one to interpret the sentence only literally, it is an anaphoric control predicate. If the idiomatic meaning is still maintained, it is a functional control predicate. For instance, using the raising verb *seem* in the English idiom (27 a) still allows an idiomatic interpretation (27 b), whereas using the control verb *try* only expresses a literal meaning (27 c).

(27) (a) The cat is out of the bag. (i.e., the secret is revealed)

(b) The cat seems to be out of the bag.

(c) #The cat tries to be out of the bag.

The same holds for the Japanese MC nouns *yotei* and *ki*. The idiomatic meaning in (28 a) is still available with the MC nominal predicate *yotei* (28 b), while the one with *ki* forces a literal interpretation (28 c).

(28) (a) *asi=ga boo=ni nar-u*
leg=NOM stick=DAT become-NPST.DECL
'The legs become sticks (i.e., exhausted)'

(b) *asi=ga boo=ni nar-u yotei=da*
leg=NOM stick=DAT become-NPST.ADN plan=COP.NPST.DECL
'The legs are going to be sticks (i.e., exhausted)'

(c) *#asi=ga boo=ni nar-u ki=da*
leg=NOM stick=DAT become-NPST.ADN plan=COP.NPST.DECL
'The legs intend to become sticks'

3.3 Semantic ambiguity

Treating MCs as instances of anaphoric/functional control predicates predicts they will interact scopally with quantifiers (Dalrymple et al. 2019, p. 556). When a raising predicate (e.g., *seem*) is used with an existentially quantified subject (e.g., *someone*), a semantic ambiguity arises as to which covers the broader scope, the raising predicate or the existential quantifier, as exemplified in (29 a). In contrast, interpretation is unique with a control predicate (29 b). This prediction is correct at least for Japanese MC nouns *ki* and *yotei*. While the scope of an existential quantifier with the functional control MC noun *yotei* has two possible interpretations (30 a), the one with the anaphoric control MC noun *ki* only allows the wide scope interpretation (30 b).

(29) (a) *Someone seemed to yawn.*

- Wide scope: $\mathsf{exist}(x, \mathsf{person}(x), \mathsf{seem}(\mathsf{yawn}(x)))$
  "There is at least one person and they seemed to yawn."
- Narrow scope: $\mathsf{seem}(\mathsf{exist}(x, \mathsf{person}(x), \mathsf{yawn}(x)))$
  "It seemed that someone yawned (but there might have been nobody)."

(b) *Someone tried to yawn.*

- Wide scope: $\mathsf{exist}(x, \mathsf{person}(x), \mathsf{try}(x, \mathsf{yawn}(x)))$
  "There is at elast one person and they tried to yawn."
- *Narrow scope

(30) (a) *dareka=ga* *ik-u* *yotei=da*
someone=NOM go-NPST.ADN plan=COP.NPST.DECL
'Someone is going to go.'

- Wide scope: $\mathsf{exist}(x, \mathsf{person}(x), \mathsf{yotei}(\mathsf{go}(x)))$
- Narrow scope: $\mathsf{yotei}(\mathsf{exist}(x, \mathsf{person}(x), \mathsf{go}(x)))$

(b) *dareka=ga* *ik-u* *ki=da*
someone=NOM go-NPST.ADN feeling=COP.NPST.DECL
'Someone intends to go.'

- Wide scope: $\mathsf{exist}(x, \mathsf{person}(x), \mathsf{ki}(x, \mathsf{go}(x)))$
- *Narrow scope

These contrasts support the idea that MCs are anaphoric and functional control constructions with a nominal predicate. Importantly, this reanalysis enables us to divide the two constructions with different syntactic structures rather than treating them all as one group. The contrasting behaviour of the two nouns also allows us to extend the typology of control and raising predicates as illustrated in Table 1.

4 Mermaid Constructions in LFG

This section presents how MCs can be accounted for in LFG. An advantage of analyzing them in LFG is that LFG's modular architecture that separates constituency and function can straightforwardly capture the intriguing structure of MCs where their main predicate is syntactically a noun but one which functionally resembles an anaphoric/functional control verb or adjective.

4.1 MCs' c-/f-structures

In MCs, the MC noun functions as the predicate of the matrix clause in a similar manner to anaphoric and functional control predicates. Therefore, the lexical entries for the Japanese MC nouns *ki* and *yotei* are as shown in (31) and (32), respectively. For simplicity, the lexical entries for control and raising stick with the assumption that raising involves structure sharing (functional control), while control involves an unpronounced PRO coindexed with the controlling argument (anaphoric control), following (Dalrymple et al. 2019, p. 545) and other lexicalist

	Verbal	Adjectival	Nominal
Control	*try*	*eager*	*ki*
Raising	*seem*	*likely*	*yotei*

Table 1: Categories (represented by an example thereof) of control and raising with respect to verbal, adjectival, and nominal predicates

theories of syntax (Landau 2013), though Asudeh (2005) does not assume PRO and explains both by structure sharing.

(31) *ki* N (↑ PRED) = 'KI⟨SUBJ, COMP⟩'
(↑ COMP SUBJ PRED) = 'PRO'

(32) *yotei* N (↑ PRED) = 'YOTEI⟨XCOMP⟩SUBJ'
(↑ SUBJ) = (↑ XCOMP SUBJ)

Upon constructing the c-structures for (1) and (2), it is assumed here that the phrase structure rules do not give a specified grammatical function for a nominal argument, because Japanese is a weakly non-configurational language and allows scrambling of nominal arguments, as exemplified in (34). To take into account this exocentricity, the phrase structure rules in (33) are assumed regarding S, where N_{MC} is a lexical category for MC nouns. The functional equation $(\uparrow \{\text{COMP} \mid \text{XCOMP}\}^* \text{GF}) = \downarrow$ for NP* allows any of the grammatical functions normally expressed by GF to be arbitrarily deeply embedded.

(33) (a) $S \rightarrow$ NP* $(\uparrow \{\text{COMP} \mid \text{XCOMP}\}^* \text{GF}) = \downarrow$ $\{VP \mid AP \mid N_{MC}P\}$ $\uparrow=\downarrow$

(b) $N_{MC}P \rightarrow$ S $(\uparrow \{\text{XCOMP} \mid \text{COMP}\}) = \downarrow$ N_{MC} $\uparrow=\downarrow$

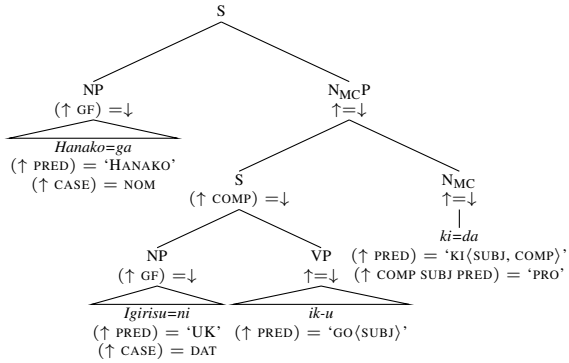
(34) *Igirisu=ni Hanako=ga ik-u yotei=da*
UK=DAT Hanako=NOM go-NPST.ADN plan=COP.NPST.DECL
'Hanako is going to go to the UK. (=(2))'

The c- and f-structures for (1) and (2) are illustrated in Figures 1 and 2. For reference, putative c- and f-structures for (1) based on the monoclausal analysis with a compound predicate are also given in Figure 3.[5] At this point, the c-structures of (1) with a *ki*-type (anaphoric control) predicate and (2) with a *yotei*-type (functional control) predicate are identical.What makes a crucial difference in these two constructions is their f-structures (or mapping from c- to f-structure).

4.2 Crosslinguistic analysis

So far, we have seen that MCs in Japanese can be successfully described by treating them as raising and control predicates. To confirm the crosslinguistic generality of this finding, this subsection illustrates various structures of MCs across languages and shows that f-structure's expressiveness can consistently account for crosslinguistic functional variation in MCs. Note that the purpose of this subsection is not

[5]Because few details on the syntactic structure inside a compound predicate are provided in Tsunoda (2020), the c- and f-structures in Figure 3 assume that compound predicates are complex predicates (Butt 1995) and follow the notations and representations given by Lowe (2016).

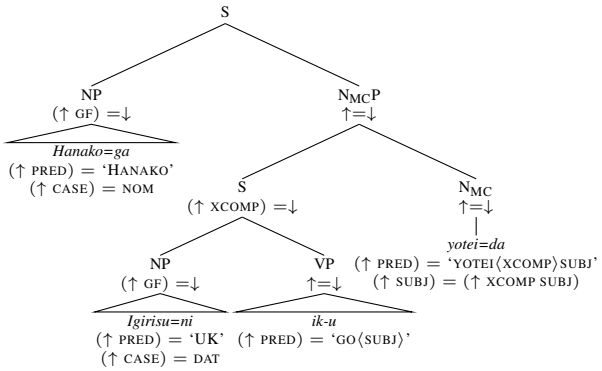


(a) c-structure

$$\begin{bmatrix} \text{PRED} & \text{'KI}\langle\text{SUBJ, COMP}\rangle\text{'} \\ \text{SUBJ} & \begin{bmatrix} \text{PRED} & \text{'HANAKO}_i\text{'} \\ \text{CASE} & \text{NOM} \end{bmatrix} \\ \text{COMP} & \begin{bmatrix} \text{PRED} & \text{'GO}\langle\text{SUBJ}\rangle\text{'} \\ \text{SUBJ} & \begin{bmatrix} \text{PRED} & \text{'PRO}_i\text{'} \end{bmatrix} \\ \text{ADJ} & \left\{ \begin{bmatrix} \text{PRED} & \text{'UK'} \\ \text{CASE} & \text{DAT} \end{bmatrix} \right\} \end{bmatrix} \\ \text{TENSE} & \text{NPST} \end{bmatrix}$$

(b) f-structure

Figure 1: c- and f-structures for (1)

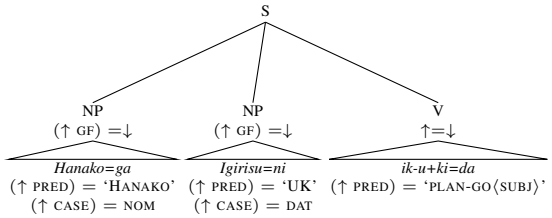


(a) c-structure

$$\begin{bmatrix} \text{PRED} & \text{'YOTEI}\langle\text{XCOMP}\rangle\text{SUBJ'} \\ \text{SUBJ} & \begin{bmatrix} \text{PRED} & \text{'HANAKO'} \\ \text{CASE} & \text{NOM} \end{bmatrix} \\ \text{XCOMP} & \begin{bmatrix} \text{PRED} & \text{'GO}\langle\text{SUBJ}\rangle\text{'} \\ \text{SUBJ} & \\ \text{ADJ} & \left\{ \begin{bmatrix} \text{PRED} & \text{'UK'} \\ \text{CASE} & \text{DAT} \end{bmatrix} \right\} \end{bmatrix} \end{bmatrix}$$

(b) f-structure

Figure 2: c- and f-structures for (2)

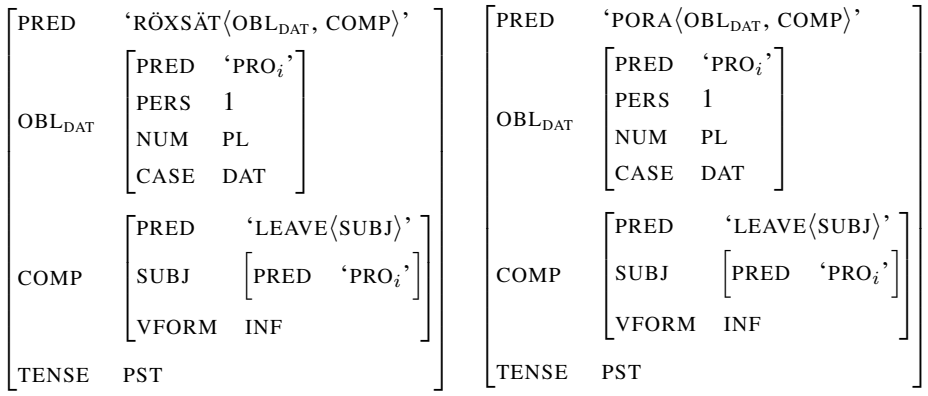


(a) c-structure

$$\begin{bmatrix} \text{PRED} & \text{'PLAN-GO}\langle\text{SUBJ}\rangle\text{'} \\ \text{SUBJ} & \begin{bmatrix} \text{PRED} & \text{'HANAKO'} \\ \text{CASE} & \text{NOM} \end{bmatrix} \\ \text{ADJ} & \left\{ \begin{bmatrix} \text{PRED} & \text{'UK'} \\ \text{CASE} & \text{DAT} \end{bmatrix} \right\} \end{bmatrix}$$

(b) f-structure

Figure 3: c- and f-structures for (1) under the compound predicate hypothesis



(a) An f-structure for (35) (b) An f-structure for (7)

Figure 4: F-structures for Tatar and Russian MCs

to provide a definitive analysis for all the MCs found across languages but to show how (non-canonical) MCs in general can also be incorporated in the LFG analysis provided above. Further language-specific discussions, such as solidly determining whether a predicate triggers anaphoric or functional control, are left for future work.

4.2.1 Dative–Infinitive MC in Tatar and Russian: *röxsät*, *pora*

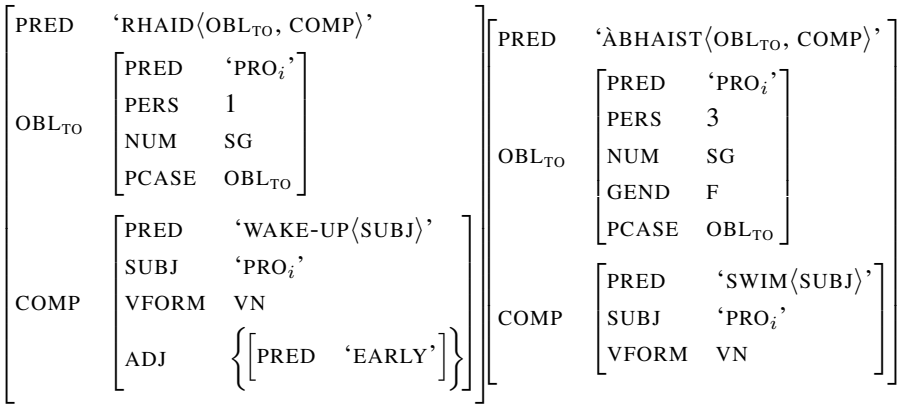
The Tatar MC noun *röxsät* “permission” expresses dynamic modality (“be allowed to”), taking an infinitive for the event predicate and a dative argument as the agent of the event, e.g., (35).

(35) Tatar
bez-gä *kit-ärgä* *röxsät* *bul-dï*
1PL-DAT leave-INF permission be-PST.3
‘We were allowed to leave.’

The structure of the Russian MC with *pora* “time”, which we already saw in (7), can be described similarly to the Tatar MC with *röxsät*, as illustrated in Figure 4. The difference in their word order is a matter of c-structures. Notice that the dative argument is assumed to be OBL_{TO} and not SUBJ; this treatment is not conclusive, as the status of their subjecthood is controversial (Bailyn 2011).

4.2.2 Dative–Verbal Noun MC in Welsh and Scottish Gaelic: *rhaid*, *àbhaist*

We saw in (5), the Welsh MC with *rhaid* “necessity”, that Welsh uses a prepositional phrase with *i* “to” to mark the responsible participant for the action, unlike the case inflections seen in Tatar and Russian. Nevertheless, the constructions can receive a similar analysis, with the prepositional phrase incorporating the PCASE



(a) An f-structure for (5)

(b) An f-structure for (6)

Figure 5: f-structures for Welsh and Scottish Gaelic MCs

feature to select for specific prepositions. This representation also holds valid for the Scottish Gaelic MC with *àbhaist*.

4.2.3 Genitive–Infinitive MC in Tatar and Tagalog: *isäp*, *plano*

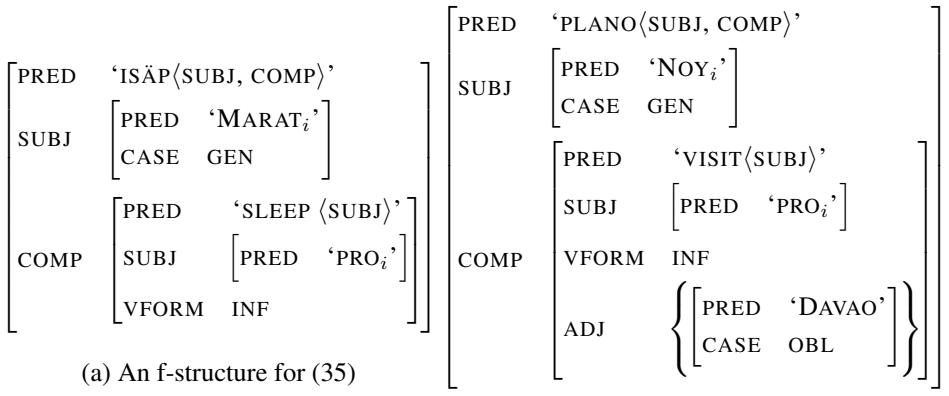
Another Tatar MC noun, *isäp* takes a genitive agent and a verb in the infinitive form (36). In addition to this, the Tagalog MC noun *plano* is also reported to take a genitive agent with the event verb expressed in the infinitive form (37) (Katagiri 2020). Their f-structures are juxtaposed in Figure 6.[6]

(36) Tatar
Marat-nïŋ joqla-rɣa isäb-e
Marat-GEN sleep-INF thought-POSS.3
‘Marat is going to sleep.’

(37) Tagalog
plano ni=Noy na b‹um›isita sa=Davao
plan GEN=Noy LK visit‹AF.INF› OBL=Davao
‘Noy plans to visit Davao.’

More interestingly, the genitive subject in the *plano* construction can appear inside the subordinate infinitival clause as in (38). Though further investigation of this construction is necessary, we can assume backward raising of the genitive subject (Wurmbrand 2015) as illustrated in Figure 7. Note that the direction of raising is only evident at the level of the c-structure.

[6]The genitives as SUBJs are a tentative treatment, as it has been an open question in LFG as to how the common properties of possessors and subjects as grammatical functions should be represented in an f-structure (Bresnan et al. 2015, p. 100). See Taguchi (2022) on the details of the Tatar genitive-infinitive MCs. See Kroeger (1991) on case and subjecthood in Tagalog.



(a) An f-structure for (35)

(b) An f-structure for (37)

Figure 6: f-structures for the MCs with Tatar *isäp* and Tagalog *plano*

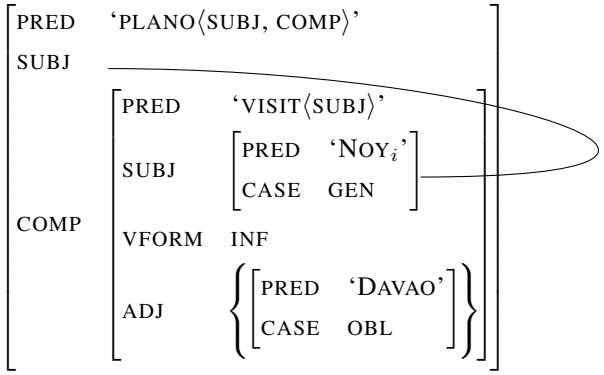


Figure 7: An illustration of backward raising in (38)

(38) Tagalog
plano=ng b‹um›isita ni=Noy sa=Davao
plan=LK visit‹AF.INF› GEN=Noy OBL=Davao
'Noy plans to visit Davao.'

4.2.4 Possessive–Finite MC in Egyptian Arabic: *šakl*

It has been reported that the Egyptian Arabic noun *šakl* causes raising (ElSadek and Sadler 2015, Camilleri and Sadler 2019). While the subordinate clause is finite, its agent's person and number are copied into the possessive suffix on *šakl* as in (22). In (22), in addition, the subject *Mary* is topicalized.

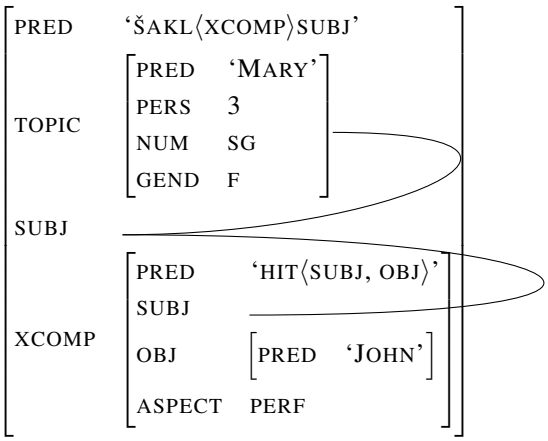


Figure 8: Raising in the Egyptian Arabic MC (22)

4.2.5 Nominative–Finite MC in Tagalog: *mukha*

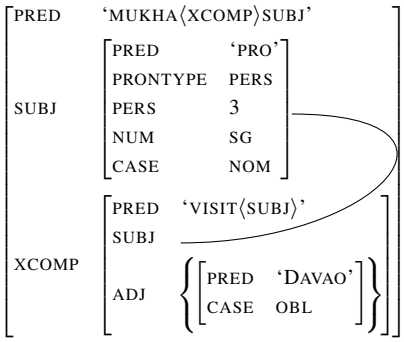
Last but not least, the Tagalog MC noun *mukha* exhibits interesting syntactic behaviors. When the subject is expressed by a pronoun, it can optionally be raised to the Wackernagel second position of the matrix clause, as contrasted in (39) and (40). Both expressions share the fact that the subject is not a semantic argument of *mukha* and that the subordinate clause is a finite clause with its own subject, i.e., COMP. The subject of COMP may optionally be raised to the matrix clause as an enclitic only if the subject is a personal pronoun. Considering the condition of subject raising, the lexical entry for *mukha* with a pronominal subject is tentatively described as MUKHA⟨XCOMP⟩SUBJ or MUKHA⟨COMP⟩ if (↑ {XCOMP | COMP} SUBJ PRONTYPE) = PERS.[7] With this assumption, the f-structures of (39) and (40) can be expressed as in Figure 9.

(39) Tagalog
mukha=siya=ng [*bibisita* *sa=Davao*]
face=3SG.NOM=LK CONT.AF:visit OBL=Davao
‘It seems that he will visit Davao.’

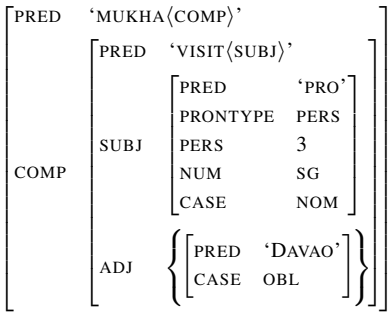
[7]Note that the *mukha* construction with a non-pronominal subject resembles adverbial modal modification by *para* “seemingly” which has a monoclausal structure (Katagiri 2020). From the rigidly lexicalist viewpoint, it is also possible to assume two different lexical entries for *mukha*:

(i) *mukha*$_1$ N (↑ PRED) = ‘MUKHA⟨XCOMP⟩SUBJ’
(↑ SUBJ) = (↑ XCOMP SUBJ)
(↑ SUBJ PRED PRONTYPE) = PERS

(ii) *mukha*$_2$ Adv (↑ PRED) = ‘SEEMINGLY’



(a) An f-structure of (39)



(b) An f-structure of (40)

Figure 9: f-structures for MCs with *mukha*

(40) Tagalog
mukha=ng [*bibisita=siya* *sa=Davao*]
face=LK CONT.AF:visit=3SG.NOM OBL=Davao
'It seems that he will visit Davao. (=39)'

5 Conclusion

This paper presented a theoretical account of the various idiosyncratic syntactic behaviors of MCs in the LFG framework. In doing so, Tsunoda's (2020) claim that MCs have a monoclausal structure with a compound predicate was rejected, and the new proposal instead argued that MCs have a biclausal structure with anaphoric/functional control triggered by a nominal predicate. Tsunoda's (2020) monoclausal hypothesis has the following problems: first, crosslinguistic data revealed it is not always the case that the (embedded) verb and the MC noun are adjacent in word order, where a compound predicate cannot be formed; second, the evidence such as the scope of negation, double temporal adverbials, honorific agreement, and the position and the case-marking of the subject in Japanese suggests that MCs are biclausal. It was then demonstrated using the syntactic diagnostics and the semantic ambiguities that MCs have identical properties to anaphoric/functional control constructions.

Section 4 outlined how Japanese MCs and those in other languages can be analyzed in the LFG framework. The modular architecture of LFG enables us to analyze MCs that involve a syntactically nominal predicate with functional similarities to attitude predicates and predicates with aspectual, modal, and evidential meanings. Finally, the anaphoric/functional control analysis was applied to MCs across languages to show the possibility of generalizing the findings. However, the analyses in this paper heavily depend on the evidence from Japanese, and it is still possible that some MCs in other languages have a monoclausal structure involving

a complex predicate or some other structure. Further analyses for each language are left for future work.

List of abbreviations

ACC: accusative; ADN: adnominal modifier form; AF: actor focus; CONT: contemplative; COP: copula; DAT: dative; DECL: declarative form; F: feminine; GEN: genitive; HON: honorification; HRS: hearsay; INF: infinitive; LK: linker; N: neuter; NOM: nominative; NPST: non-past; OBL: oblique; PASS: passive; PFV: perfective; PL: plural; POSS: possessive; PST: past; SG: singular; TOP: topic; VN: verbal noun.

References

Asudeh, Ash. 2005. Control and semantic resource sensitivity. *Journal of Linguistics* 41(3), 465–511.

Bailyn, John Frederick. 2011. *The syntax of Russian*. Cambridge University Press.

Bresnan, Joan, Asudeh, Ash, Toivonen, Ida and Wechsler, Stephen. 2015. *Lexical-functional syntax*. Wiley, second edition.

Butt, Miriam. 1995. *The structure of complex predicates in Urdu*. CSLI Publications.

Butt, Miriam. 2014. Control vs. complex predication: identifying non-finite complements. *Natural Language & Linguistic Theory* 32(1), 165–190.

Camilleri, Maris and Sadler, Louisa. 2019. A non-canonical diachronic formation of raising predicates. In Miriam Butt, Tracy Holloway King and Ida Toivonen (eds.), *Proceedings of the LFG19 Conference*, CSLI Publications.

Dalrymple, Mary, Lowe, John J. and Mycock, Louise. 2019. *The Oxford reference guide to Lexical Functional Grammar*. Oxford University Press.

ElSadek, Shaimaa and Sadler, Louisa. 2015. Egyptian Arabic perceptual reports. In Miriam Butt and Tracy Holloway King (eds.), *Proceedings of the LFG15 Conference*, CSLI Publications.

Katagiri, Masumi. 2020. Tagalog. In Tasaku Tsunoda (ed.), *Mermaid Construction: a compound-predicate construction with biclausal appearance*, Comparative Handbooks of Linguistics, pages 781–816, Mouton de Gruyter.

Kawachi, Kazuhiro. 2020. Sidaama. In Tasaku Tsunoda (ed.), *Mermaid Construction: a compound-predicate construction with biclausal appearance*, Comparative Handbooks of Linguistics, pages 679–736, Mouton de Gruyter.

Kroeger, Paul R. 1991. *Phrase structure and grammatical relations in Tagalog*. Ph.D. thesis, Stanford University.

Landau, Idan. 2013. *Control in generative grammar: a research companion*. Cambridge University Press.

Li, Chao. 2004. On verbal nouns in Celtic languages. In *Proceedings of the Harvard Celtic Colloquium*, volume 24/25, pages 163–194.

Lowe, John J. 2016. Complex predicates: an LFG+Glue analysis. *Journal of Language Modelling* 3(2), 413–462.

Matsumoto, Yo. 1996. *Complex predicates in Japanese: a syntactic and semantic study of the notion 'word'*. CSLI Publications.

Matsumoto, Yoshiko. 1988. Semantics and pragmatics of noun-modifying constructions in Japanese. In *Proceedings of the Fourteenth Annual Meeting of the Berkeley Linguistics Society*, pages 166–175.

Polinsky, Maria. 2013. Raising and control. In Marcel Den Dikken (ed.), *The Cambridge handbook of generative syntax*, Cambridge University Press.

Ramchand, Gillian Catriona. 1993. Verbal nouns and event structure in Scottish Gaelic. In *Proceedings of the Semantics and Linguistic Theory conference (SALT III)*.

Shimoji, Michinori. 2020. Irabu Ryukyuan. In Tasaku Tsunoda (ed.), *Mermaid Construction: a compound-predicate construction with biclausal appearance*, Comparative Handbooks of Linguistics, pages 233–282, Mouton de Gruyter.

Taguchi, Chihiro. 2022. *Deconstructing Mermaid Constructions: from control and raising in Tatar*. Master's Thesis, University of Edinburgh.

Tsunoda, Tasaku. 2011. Ningyo koubun: nihongogaku kara ippan gengogaku he no kouken. *NINJAL Research Papers* 1, 53–75.

Tsunoda, Tasaku. 2012. Ningyo koubun to meisi no bunpouka. *NINJAL Project Review* 7, 3–11.

Tsunoda, Tasaku (ed.). 2020. *Mermaid Construction: a compound-predicate construction with biclausal appearance*. Comparative Handbooks of Linguistics, Mouton de Gruyter.

Wurmbrand, Susi. 2015. On forward and backward raising. In Ulrike Steindl, Thomas Borer, Huilin Fang, Alfredo García Pardo, Peter Guekguezian, Brian Hsu, Charlie O'Hara, and Iris Chuoying Ouyang (eds.), *Proceedings of the 32nd West Coast Conference on Formal Linguistics*, pages 342–351, Cascadilla Proceedings Project.

Zaenen, Annie and Dalrymple, Mary. 1995. "Polymorphic" causatives: complex predicates in French. *AAAI Technical Report* SS-95-01, 180–185.