문학석사학위논문

Embedded Topicalization in Korean Factive Complement Clauses: An Experimental Approach

한국어 사실성 구문에서의 화제화에 대하여:
실험적 접근

2018년 8월

서울대학교 대학원
언어학과 언어학전공
주 은 선

Embedded Topicalization in Korean Factive Complement Clauses: An Experimental Approach

지도교수 고 희 정

이 논문을 문학석사 학위논문으로 제출함
2018년 6월

서울대학교 대학원
언어학과 언어학전공
주 은 선

주은선의 문학석사 학위논문을 인준함
2018년 7월

위 원 장 남 승호 (인)

부위원장 고 희정 (인)

위 원 정 한빛 (인)

Abstract

Embedded Topicalization in Korean Factive Complement Clauses: An Experimental Approach

Eunsun Jou

Department of Linguistics

The Graduate School

Seoul National University

Complement clauses of factive predicates are considered to be structurally reduced compared to their non-factive counterparts. Semantically, they are said to lack assertion – their propositional content is not delivered from the viewpoint of the speaker but simply presupposed, or presented as is (Hooper and Thompson 1973). Syntactically, it is said to disallow root transformations inside of them. Root transformations are a group of syntactic movements including negative constituent preposing, left/right-dislocation and topicalization. They are only available in root clauses and "root-like" embedded environments (Emonds 1976, Hooper and Thompson 1973).

The root transformation that this study focuses on is topicalization. Topicalization

is one of the root transformations that have been thoroughly debated since the original inception of Emonds (1976). When a phrase is topicalized, it moves to the clause-initial position and receives the sentential topic interpretation. Factive complement clauses do not license this movement inside itself.

To explain why factive complement clauses do not allow topicalization, one must take into consideration the precise structure of factive clauses as well as the nature of topicalization. This study examines previous theories on this issue and recognizes two different approaches that account for lack of topicalization inside factive complement clauses: the operator approach and the clause-size approach.

The operator approach is more semantically motivated. It is inspired by the intuition of Melvold (1991) that a factive complement clause is a definite description of an event. Watanabe (1993) and Haegeman and Ürögdi (2010) advocate this approach and aim to explain lack of topicalization as a result of competition for the same landing site or of an featural intervention effect.

The clause-size approach appeals to structural differences between factive and non-factive complement clauses. Mostly laid out in the early works of Haegeman (2004, 2006), it views factive complement clauses as structurally smaller than non-factive complement clauses. Her view connects to the classic arguments of Hooper and Thompson (1973) that factive clauses lack assertion. What is exclusively projected in non-factive complement clauses encode information such as viewpoint and

assertion (Rizzi 1997, Haegeman 2004, De Cuba 2007). Under this view, topicalization is unavailable in factive complement clauses because it requires one of the functional projections that are present in non-factive complement clauses but missing in the factive ones.

However, the situation is more complicated in the case of Korean. In Korean, factive predicates can select two different forms of complement clauses. The "large" form projects a series of functional structures including the sentence final particle that appears in non-factive complement clauses and matrix clauses. The "small" form only projects a part of the functional structures projected in its larger counterpart. According to the logic of the operator approach, topicalization would be unavailable in both cases since the factive operator at the clause edge would hinder topic licensing no matter the size of the clause. On the other hand, the clause-size approach expects topicalization to be available in the larger clauses since it projects the necessary structure to license topic.

To the best of my knowledge, these two approaches were never applied to analyze Korean data in previous studies; therefore I judged it would be beneficial to attempt analyzing Korean data with these frameworks and see whether their predictions actually bear out. An acceptability judgment experiment was carried out in order to empirically verify the predictions of each approach. The results demonstrate that both the operator approach and the clause-size approach make correct predictions for fac-

tive complement clauses in Korean. A follow-up experiment was designed to verify whether the clause-size approach can be extended to explain the lack of topicalization inside islands. The results provide partial support for the clause-size approach.

This study analyzes a new set of data with existing theories, delineates the predictions of each theory, and verifies the prediction by means of a formal experiment. Based on the results, it calls for a reconciliation of the two approaches in explaining factivity in Korean. Emphasizing the explanatory adequacy of the clause-size approach in factive complement clauses as well as islands, it also advocates a split-CP framework approach to various phenomena at the syntax-pragmatics interface of Korean.

Keywords: factivity, embedded clauses, topicalization, experimental syntax, syntax-pragmatics interface

Student Number: 2016-20061

Contents

1 Introduction

This study focuses on the lack of topicalization inside factive complement clauses. There are mainly two different theories that address this issue: the operator approach and the clause-size approach. This study is an attempt to experimentally verify whether the predictions of the two approaches hold for Korean. Korean is a language that, unlike most other languages, possesses two types of factive clauses – possibly of different clausal size. Previous theories have not taken the possibility of a two-way system of factivity into consideration. This study applies the theories on Korean data and explores what they predict about the grammaticality of embedded topicalization inside factive clauses. It verifies whether the predictions are borne out via an acceptability judgment experiment. The results suggest that both approaches are needed to adequately explain the data on factive clauses.

1.1 Root transformations, topic and factivity

Since the influential observation of Emonds (1969, 1976), the notion of "rootness" and root transformation has been an important theme in syntax. It is also called *main clause phenomena* in some studies (Haegeman 2006, Haegeman and Ürögdi 2010). Root transformations are, as their name suggests, transformations that are available in root clause environments. It is usually related to affective or pragmatic effects such as negative constituent preposing, topicalization, left dislocation, and subject-auxiliary inversion. Hooper and Thompson (1973) connect this pragmatic aspect of root trans-

formation to the notion of assertion, while Haegeman (2006) labels it speaker deixis or speaker anchoring.

An interesting point is that root transformations are not strictly limited to matrix clauses. Some embedded environments also allow root transformations within themselves. In (1), negative constituent preposing occurs inside the complement clause of the verb *exclaimed*.

(1) I exclaimed that never in my life had I seen such a crowd.

(Hooper and Thompson 1973: 474)

Examples like this where root transformations occur inside embedded clauses indicate that some embedded environments are "root-like" enough to allow root transformations inside of themselves. They involve some form of assertion or speaker anchoring just like matrix clauses. Various studies have attempted to distinguish what kinds of embedded clauses are root-like and what kinds are not. Successful investigation into this area will reveal much about the clause edge of embedded structures as well as inform us about phenomena at the syntax-pragmatics interface.

One type of environment that is known to disallow root transformations is complement clauses of factive predicates (henceforth *factive clauses*). Factive predicates are a subset of predicates that select a clausal complement. They display a unique semantic behavior: their complement clauses require a factive presupposition. That is, when the predicate and its complement factive clause is uttered, the complement

must be accepted as true by the speaker. Otherwise the utterance is infelicitous. For example, a speaker who utters (2) would believe the complement clause *that I didn't attend the concert* to be true.

(2) I regret that I didn't attend the concert.

(Hooper and Thompson 1973:479)

Since Hooper and Thompson (1973) define rootness as involving assertion, it is no surprise that they attempt to account for the lack of root transformations in factive clauses with their lack of assertion. Indeed, other early studies have also understood presupposition as a lack of assertion (Kiparsky and Kiparsky 1970, Stalnaker 1975). Hooper and Thompson's approach can be understood within such a context, and the presupposition-assertion contrast still remains relevant, which is the theme underlying the research question of related studies (Melvold 1991, Haegeman 2006, De Cuba 2007 among others).

Among the various root transformations discussed in the literature, topicalization is one of the most thoroughly discussed phenomena. Topic is a fundamental concept of information structure and discourse, functioning as a link to older discourse and crucial in forming a felicitous utterance (Reinhart 1981, Lambrecht 1994, Vallduví 1990). Cross-linguistically, speakers employ a variety of linguistic strategies to express topic including morphological, syntactic, and prosodic elements (Gundel 1988). Topicalization is a strategy to express topichood with a syntactic device: movement

to the head of the clause. In (3), the noun *apples* undergoes topicalization.

(3) Apples$_{i}$, I gave to John t_{i}.

Because this study aims to test some movement-related consequences such as featural intervention, it focuses on topicalization and its interaction with factive clauses.

Syntactic theories that account for the lack of topicalization inside factive clauses can be classified into two groups. The operator approach postulates an operator at the periphery of the factive clause and attributes the lack of topic licensing to this operator (Watanabe 1993, Haegeman and Ürögdi 2010). On the other hand, the clause-size approach does not employ an operator but instead stipulates a hierarchy of functional heads at the clause edge (Haegeman 2004, 2006). Factive clauses lack some of the functional heads that non-factive embedded clauses project; in this approach, topic is not licensed because factive clauses do not include the appropriate head that licenses them.

1.2 An experimental approach

The aim of this study is to evaluate the two approaches against newly collected data in Korean and to determine whether the two serve as a reliable model for them. Unlike most other languages, there are two different forms of factive clauses in Korean.

(4) a. Hani-ka [Minho-ka cwungkwuke-lul calha-nun kes]-ul
Hani-NOM Minho-NOM Chinese-ACC do.well-ADN KES-ACC
nollaweha-yss-ta.
be.surprised-PST-SFP

"Hani was surprised that Minho speaks Chinese well."

b. Hani-ka [Minho-ka cwungkwukelul calha-n-ta-nun
Hani-NOM Minho-NOM Chinese-ACC do.well-PRES-SFP-ADN
kes]-ul nollaweha-yss-ta.
KES-ACC be.surprised-PST-SFP

"Hani was surprised that Minho speaks Chinese well."

Korean serves as a testing ground for the two approaches because they provide different predictions on the same pair of structures. The operator approach expects embedded topicalization to lead to ungrammaticality in both (4b) and (4a). The clause-size approach, on the other hand, predicts that topicalization would be possible in (4b) but not in (4a).

However, judgment on topic and factivity is quite subtle, and it is difficult for researchers to compare the two approaches solely based on their own judgment. An experimental approach allows to systematically factorize the phenomenon of interest (topicalization inside factive clauses) and to collect a robust sample of judgments from multiple speakers. Hence an acceptability judgment experiment was designed. Participants were presented a stimulus sentence and asked to mark its grammaticality on a Likert scale. The results suggest that the prediction of both approaches are borne out, calling for a reconciliation of the two theories.

A second experiment was designed as an attempt to verify whether the clause-

size approach corroborated by the results of the first experiment can also be applied to Korean island constructions. It was designed and conducted in a manner parallel to the first experiment. The results suggest that explanations pertaining to clause size are applicable to islands as well as factive clauses.

1.3 Organization of the thesis

The thesis is organized as follows. Chapter 2 provides an overview of previous studies on topic, the syntax of topicalization, and factive clauses. Based on this overview, it compares the two approaches on topicalization inside factive clauses. It is also demonstrated that according to their logic, their predictions differ from each other for Korean data. Chapter 3 explains the design of the main experiment and reports on the results. Chapter 4 presents the design and results of the follow-up experiment. Chapter 5 elaborates on four discussion points related to interpretation of the results. Chapter 6 is the conclusion.

2 Previous studies on topic and factive clauses

2.1 The syntax and semantics of topic

2.1.1 The definition of sentential topic

Topic is one of the most important concepts of information structure, yet various studies have suggested vastly different definitions for it. Topic "limits the applicability of the main predication to a certain restricted domain" (Chafe 1976), "set[s] a framework in naming what the sentence is about" (Li and Thompson 1981), or is a means to "organize, or classify the information exchanged in linguistic communication" (Reinhart 1981).

It should be clarified which of the various concepts of topic this study adopts. First, the study exclusively looks into sentence-level topic as opposed to discourse-level topic. The latter notion roughly corresponds to the subject matter of the topic or *question under discussion* (QUD), and is covered in the works of the Prague School and of Halliday (1967) and Halliday and Matthiessen (2004).[1]

The sentence-level topic basically pertains to "what the sentence is about." Within the sentence-level discussions of topic, there are again multiple approaches to the concept. Here I concentrate on *sentential topic* in the sense of Reinhart (1981). For Reinhart, a context set is not simply a set of mutually unrelated propositions but an

[1] See Roberts (2011) for a comparative overview of discourse-level and sentence-level topic, with relevant literature.

organized structure where discourse participants systematically update information under relevant entries – possibly like a dictionary or encyclopedia. A newly uttered sentence conveys some new information to be updated in the context set. Sentential topic indicates the entry under which the new information is to be updated. It is similar to Gundel's (1985, 1988) *syntactic topic* since both must be represented by a linguistic constituent within the sentence, typically at the clause-initial position.[2] Also relevant is Vallduví's (1990) *link*, which is intended to be an extension of Reinhart's sentential topic. A link is a strictly clause-initial element that performs the task of "linking up with the object of thought" (Vallduví 1990: 59). This study follows Reinhart's original terminology and calls the relevant concept *sentential topic*.

What characterizes sentential topic is its givenness, which has been roughly explained as being predictable or being salient in the discourse. (See Prince (1981) for an overview on givenness.) In an effort to rigorously define givenness, Gundel (1985, 1988) differentiates two different dimensions of givenness: referential givenness and relational givenness. Referential givenness is related to the speaker being cognitively aware of some entity. If an entity is referentially given, there exists some representation in the speaker's mind that corresponds to that entity (Gundel and Fretheim 2004). Philosophical concepts such as existential presupposition (Strawson 1964), referentiality and assumed familiarity (Prince 1981) are closely related with referen-

[2]While she agrees with the general spirit of Reinhart (1981), Gundel (1985) disagrees with Reinhart in that the sentence-level topic need not *necessarily* be represented in the syntax. Her *syntactic topic* is a subset of the wider concept *pragmatic topic*, which is defined at the sentence level ("what the sentence is about") but need not be represented linguistically.

tial givenness.

Relational givenness, on the other hand, is a more linguistic term since it denotes a certain part of a clause or sentence. As its name suggests, it is defined as a relation between two linguistic entities. X, a part of some clause, is given *in relation to* Y, another part of the clause, when Y is new information predicated about X (Gundel and Fretheim 2004). The relationally new Y is the foregrounded material, or the part that is newly asserted in the discourse.

Referential and relational givenness are clearly different terms, although they are closely related to one another. Consider the conversation in (5), where the uppercase word (PORK) indicates that it is given prominent pitch accent.

(5) A. Did you order the chicken or the pork?

B: It was the PORK I ordered. (Gundel and Fretheim 2004:177)

As Gundel and Fretheim explain, PORK in (5B) is referentially given: the speaker is aware of its existence since it is mentioned in the previous discourse (5A) and also since he or she had ordered it (for example at a restaurant). However, it is relationally new in contrast to the given element: *what the speaker of (5B) ordered*. It is the new information provided in response to the request of (5A).

As for sentential topic, it is both referentially and relationally given (Gundel 1985). It is relationally given by definition: it is the "old" part of the sentence that functions as a link or entry under which the relationally new part can be integrated

into the context set. In other words, it is given in relation to the new information part for which it serves as the entry. Its referential givenness is a precondition that follows naturally from its relational givenness. In order to employ some concept as a link to older discourse, the speaker must be cognitively aware of that concept.

It should also be made clear that the scope of this study excludes contrastive topic. Unlike non-contrastive (sentential) topic, the use of contrastive topic suggests the existence of an alternative set (Büring 2003, Lee 2003b). This effect is achieved by providing only a partial answer to a question that involves a whole set of entities or concepts. Consider Lee's (2003b) example below. (Intonation marking from the original example is omitted.)

(6) **Q**: Ne ton iss-ni?
you money have-Q
"Do you have money?"

A: Na tongcen-un iss-e.
I coins-CT have-DEC
"I have coins$_{CT}$, (but not bills.)" (Lee 2003b:2)

Here, speaker Q asks a question regarding the entire set of money types, namely {bills, coins}. However, speaker A only provides information about a subset: {coins}. The mismatch between the two evokes implicature about the unmentioned element(s) of the entire set: that speaker A does not have bills, does not have information about bills, or wishes not to talk about bills. Contrastive topic and sentential topic display differences in syntactic and semantic behavior, and researchers have considered

them to be two qualitatively different concepts (Frascarelli and Hinterhölzl 2007, Vermeulen 2009, 2012). The distinction is especially important for Korean, where both types of topic are marked with the suffix *nun* (Lee 2003b, Jun 2006, 2015). This study limits the scope of its discussion to the use of *nun* as the sentential topic marker.

To summarize, this study focuses on non-contrastive sentential topic: the relationally (and referentially) given part of the proposition which is linguistically represented as a clause-initial constituent and is interpreted as the entry under which the information conveyed by the sentence is to be integrated into the context set. In Korean, sentential topic is marked by the topic marker *nun* in addition to its clause-initial position. From here on the term "topic" refers to this non-contrastive sentential topic.

2.1.2 The syntax of topic: subtypes of the topic construction

Languages employ various morphological, intonational and syntactic devices to express the topic status of some phrase. In terms of syntax, topic is generally placed at the head of the clause. However, there are at least three variants of this phenomenon. Mandarin Chinese demonstrates a wealth of information structural syntactic phenomena; a review of topic constructions in the language provides an adequate overview of the variants (Shyu 1995, 2014, Shi 2000, Huang et al. 2009, Badan and Del Gobbo 2010).

First, there are cases where topic is semantically linked to a theta position within

the same clause. *Topicalization* indicates cases where a topic phrase receives a theta role from the verb and then moves to a higher position, leaving a gap as in (7).

(7) Zhangsan$_i$, wo kanjian t_i le.
Zhangsan I saw MOD
"Zhangsan, I saw." (Badan and Del Gobbo 2010)

Island effects and binding restrictions in this type of topic construction are evidence that movement has indeed taken place. In (8a-c), topicalization out of a complex NP, left branch or adjunct leads to ungrammaticality.

(8) a. Complex NP island

* Lisi$_i$, wo renshi [henduo [[e$_i$ xihuan] de] ren].
Lisi I know many like DE person
"Lisi$_i$, I know many people who e$_i$ likes."

b. Left Branch Condition

* Zhangsan$_i$, wo kanjian-le [e$_i$ baba].
Zhangsan I see-LE father
"Zhangsan$_i$, I saw [his$_i$] father.

c. Adjunct Condition

* Lisi$_i$, zhe-jian shi [gen e$_i$ mei lai] mei you guanxi.
Lisi this-CL matter with not come not have relation
"Lisi$_i$, this matter is not related to [his$_i$] not having come."

(Huang et al. 2009, Shyu 2014)

Similarly, the ungrammaticality of (9b) also suggests movement. It is a typical ex-

ample of Condition C violation where the R-expression *Zhangsan* is bound by the pronoun *ta* (Chomsky 1981). Had *Zhangsan* not been placed there via movement, as in (9a), there would be no reason for ungrammaticality.

(9) a. Zhangsan$_i$, ta$_i$ zou-le.
Zhangsan he leave-LE
"Zhangsan$_i$, he$_i$ left."

b. * Zhangsan$_i$, ta$_i$ bu renshi t_i.
Zhangsan he not know
"Zhangsan$_i$, he$_i$ doesn't know." (Huang et al. 2009)

Another case of topic construction related to a theta position is *left-dislocation*, where topic is co-referential with an overt resumptive pronoun in the theta position of the clause.

(10) Zhangsan$_i$, wo kanjian ta$_i$ le.
Zhangsan I saw him MOD
"Zhangsan, I saw." (Badan and Del Gobbo 2010)

Left-dislocation is not considered to involve movement, as it does not show island effects – compare (11) with (8a). Shyu (1995, 2014) suggests that they are base-generated in their position.

(11) Lisi$_i$, wo renshi [henduo [[ta$_i$ xihuan] de] ren].
Lisi, I know many he like DE person
"Lisi$_i$, I know many people who he$_i$ likes." (Shyu 2014)

The asymmetry between topicalization and left dislocation in terms of island effects also holds in English, as pointed out in Lasnik and Saito (1992).

(12) a. This book, I accept the argument that John should read it.

b. * This book, I accept the argument that John should read *t*.

Unlike topicalization and left dislocation, *hanging topic* is an instance of topic that is not linked to a theta position. Instead of being in a thematic relation with the predicate, it forms a looser "aboutness relation" with the rest of the clause. Shyu (2014) and Huang et al. (2009) consider hanging topics to be base-generated since it can be coreferential with an element inside an island as in (13b).

(13) a. Shuiguo, wo zui xihuan xiangjiao.
fruit I most like banana
"(As for) fruits, I like bananas most."

b. Shuiguo, wo zui xihuan [[bu pa chi xiangjiao de] ren].
fruit I most like not afraid eat banana DE person
"(As for) fruits, I like the most people who are not afraid to eat bananas."

(Huang et al. 2009)

2.1.3 Topic constructions in Korean

It has been mentioned that Korean marks topic with the suffix *nun*. It has been debated whether *nun* is the only topic marker and whether the function of *nun* is to truly mark topic. However, it is generally agreed that at least for a non-accentuated, clause-initial

phrase accompanied by *nun*, one can safely assume that it is indeed sentential topic (Jun 2009, Vermeulen 2009, 2012, Kim 2015).

Thus it can be said that Korean uses syntactic (clause-initialness) as well as morphological (*nun*) strategies to indicate topic. Korean also shows the three types of topic constructions that were previously introduced for Chinese.

(14) a. Topicalization

Sakwa$_i$-nun Hani-ka t_i mek-ess-ta.
apple-TOP Hani-NOM eat-PST-SFP

"(The) apple$_i$, Hani ate t_i."

b. Left-dislocation

Seoul$_i$-un Toli-ka taumtal-ey keki-ey ka-lyeko
Seoul-TOP Toli-NOM next.month-LOC there-LOC go-CONNECT
saynggakha-ko iss-ta.
think-Comp be-SFP

"Seoul$_i$, Toli is thinking of going there$_i$ next month." (Moon 1994)

c. Hanging topic

Kwail-un sakwa-ka mas-iss-ta.
fruit-TOP apple-NOM taste-be-SFP

"(As for) fruits, apples taste best."

Of the three subtypes of topic, this study focuses on topicalization (14a), which is the only case that involves movement. Subsections 2.3.2-2.3.3 will compare two different approaches on embedded topicalization inside factive clauses: the operator ap-

proach and the clause-size approach. A version of the operator approach involves featural relativized minimality and intervention. Since relativized minimality is a locality constraint that limits the range of movement, one needs a syntactic operation that involves movement in order to effectively test and compare the approach. Topicalization is such an operation and thus will be the target of this study.

2.2 Landing site of topicalization

If topicalization involves movement, a question that follows naturally is where its target position is. Since the analysis of Higgins (1973) that topicalization is substitution of Comp, there have been largely two competing ideas about the landing site of topic movement: CP-recursion and IP-adjunction. While earlier accounts weighed towards IP-adjunction, a majority of studies since Authier (1992) locate topicalization in the CP area. The remainder of this section is as follows. Subsection 2.2.1 briefly reviews the original analysis of topicalization laid out in Chomsky (1977). Subsection 2.2.2 provides an overview of the supporting arguments for the IP-adjunction and CP-recursion approach. Lastly, Subsection 2.2.3 presents a third, more recent viewpoint: topic within the split CP framework of Rizzi (1997).

2.2.1 The classic analysis of Chomsky (1977)

Chomsky's analysis of topicalization is inspired from his analysis of *wh*-movement from the same article. Here, *wh*-questions and relative clauses are considered to be

derived by the same process: the *wh*-element, which is understood as a type of quantifier, moves to the COMP position below S′ ("move *wh*-phrase into COMP").

The reason why Chomsky applies to topicalization the same analysis as *wh*-movement is because the former displays the four general characteristics of the *wh*-movement rule. The relevant characteristics are listed in (15), and (16) shows how topicalization displays the corresponding characteristics.

(15) a. It leaves a gap.

b. Where there is a bridge, there is an apparent violation of subjacency, PIC, and SSC.[3]

c. It observes CNPC.

d. It observes *wh*-island constraints. (Chomsky 1977:86)

(16) a. *this book, I really like*

b. *this book, I asked Bill to get his students to read*

c. * *this book, I accept the argument that John should read*

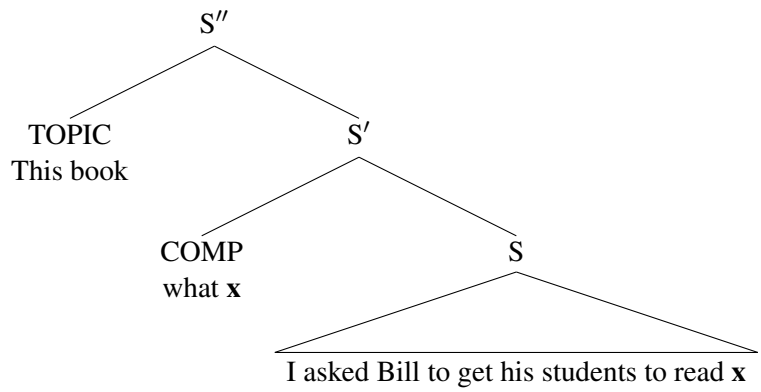
d. * *this book, I wonder who read* (Chomsky 1977:91)

Specifically, topicalization is the formation of a structure where there is some base-generated phrase at the TOPIC position above S′, and the rest of the clause (S′) is an "open sentence" that says something about the entity referred to by the structure in

[3]Here PIC stands for propositional-island conditions, and SSC for specified subject condition.

TOPIC. Chomsky admits that the S′ structure resembles the *wh*-movement, especially the free relative construction. The syntactic analysis is accordingly similar as well. The structure of a sample sentence is schematized in (17).

(17) This book, I asked Bill to get his students to read.



(Chomsky 1977:91, his (70) and (72))

Wh-movement occurs as in a regular relative clause and a bound variable x is introduced. Later, the *wh*-element is deleted under the obligatory *wh*-deletion rule, making the variable unbound and turning S′ into an open sentence. The open sentence predicates over the topic, enabling interpretation in the semantics. This derivation is reminiscent of *wh*-operator movement.

However, this analysis is not without problems. First of all, Chomsky's analy-

[3]In what Chomsky calls left-dislocation structure, TOPIC can consist of not only bare DPs but also specific expressions that explicitly mark aboutness such as *as for X* or *as far as X is concerned*. This is different from what has been discussed in section 2.1.2 where only the bare DP form was discussed.

sis expects topicalization to be unable to co-occur with *wh*-movement because both operations involve movement to COMP. Hence the ungrammaticality in (18), where topicalization occurs inside a *wh*-question and a relative clause.

(18) a. * To whom the books did John give away?

b. * the boy to whom the books John gave away

(Chomsky 1977:92)

However, Baltin (1982) points out that it is possible to construct a sentence where *wh*-movement and topicalization co-occur.

(19) He's a man to whom liberty we could never grant.

(Baltin 1982:17)

In (19), *wh*-movement of *to whom* and topicalization of *liberty* co-occur. But Baltin states that (19) is acceptable for some speakers. Since both Chomsky and Baltin assume, following Higgins (1973), that COMP cannot be doubly filled, the grammaticality of (19) is problematic for Chomsky. This observation led Baltin to come up with an alternative theory on the landing site of topicalization.

2.2.2 The IP-adjunction approach versus the CP-recursion approach

Baltin (1982) employs the analysis of Maling and Zaenen (1977) where they analyze Icelandic topicalization as adjunction to IP (or S in his terms), and applies it to English. COMP cannot be doubly filled and topic comes lower than *wh*-movement (19).

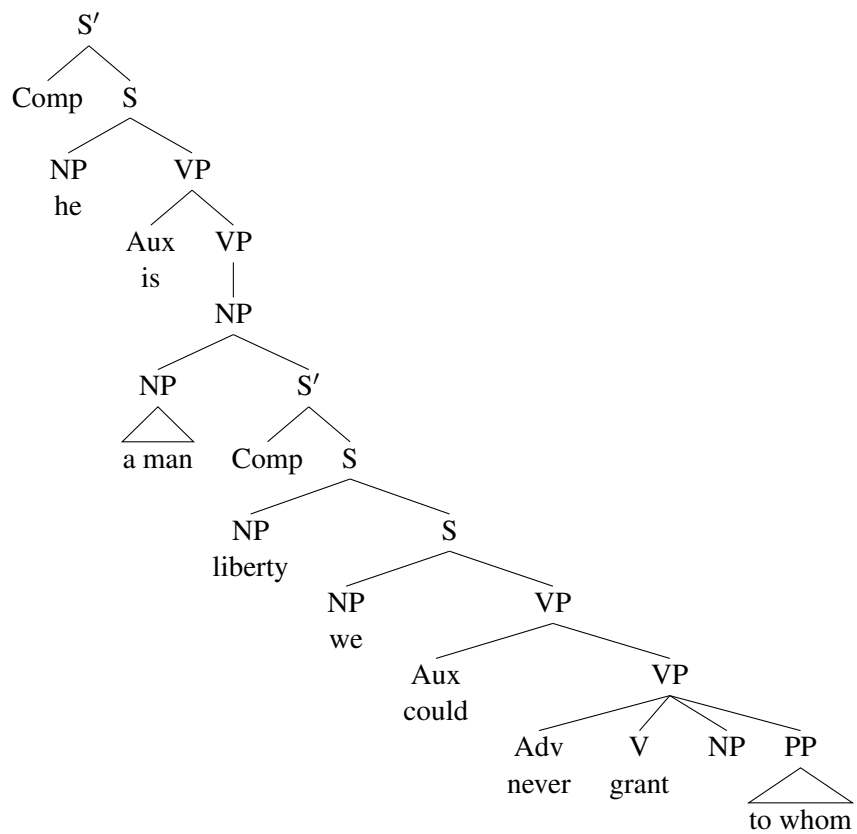
Given these facts, it is a plausible assumption to place the *wh*-element in COMP and to place the topic as an adjunct to S. Topic is also lower than COMP itself for that matter (20), which seems to be even clearer evidence that topicalization is adjunction to S.

(20) a. John says that Sue, Bill doesn't like.

b. * John says Sue, that Bill doesn't like. (Authier 1992)

Baltin's analysis of (19) is demonstrated below.

(21)



The topicalized NP *liberty* is adjoined to S, and the *wh*-phrase *to whom* would raise to COMP without any problem.

Baltin's view is shared by Lasnik and Saito (1992), especially in the context of embedded topicalization. They focus on Baltin's observation that left-dislocation is unavailable in embedded environments. They provide additional evidence in (22-23) that embedded left-dislocation is ungrammatical.

(22) a. * He's a man to whom as for liberty, we could never grant it.

b. * He's a man to whom liberty, we could never grant it.

(23) a. that this solution, I proposed last year is widely known

b. * that this solution, I proposed it last year is widely known

(Lasnik and Saito 1992:77)

They claim that the unavailability of left-dislocation in embedded environments provides further support for Baltin's theory. Since left-dislocation only involves base-generated structures in TOPIC, they claim that their unavailability in embedded environments is simply an indication that TOPIC is inexistent in embedded environments. If this is true, then Baltin's view adequately explains the grammaticality of embedded topicalization such as (23a): since there is no TOPIC structure in embedded sentences, the only position available for topicalized elements is adjunction to IP.

Authier (1992) effectively argues against the IP approach by appealing to the

similarity between topicalization and negative inversion (negative constraint preposing in Emonds's (1976) terms). If negative inversion is in the lower [Spec, CP] of the recursive CP layer (to be explained below), and topicalization targets the same position as negative inversion, topic would also be located at CP.

Below is an example of embedded negative inversion (24a) and embedded topicalization (20b, repeated here as 24b). Authier's intuition is that *never in her life* in (24a) and *Sue* in (24b) are located in the same position.

(24) a. Mary kept saying that never in her life had she seen such a thing.

b. John says that Sue, Bill doesn't like.

(Authier 1992:329)

Authier follows previous analyses of Koopman (1983) and Chomsky (1986) that negative inversion is raising of the auxiliary from I to C. Under this assumption, the instances of embedded negative inversion such as (24a) call for at least two C positions: a C that houses the raised auxiliary *had*, and another C that houses the complementizer *that*. As for the phrases *never in her life* and *Sue*, a natural candidate for their position is the specifier of the lower C *had*.

If topicalization is indeed structurally similar to negative inversion, it can also be explained as occurring in [Spec, CP]. While he admits that there is no direct evidence that negative inversion and topicalization target the same position, Authier presents data from Hooper and Thompson (1973) as indirect evidence. Their data suggest that

both negative inversion and topicalization are only available in environments where deletion of the complementizer *that* is acceptable. For example, they both cannot appear inside CP complements of nouns (25) nor inside CP subjects (26). They are both environments that disallow *that*-deletion (Stowell 1981).

(25) Negative inversion (a) and topicalization (b) inside CP complements of nouns

a. * The fact that never has he had to borrow money makes him very proud.

b. * The fact that Bill, Mary likes makes John very jealous.

(26) Negative inversion (a) and topicalization (b) inside CP subjects

a. * That never in his life has he had to borrow money is true.

b. * That this book, Mary read thoroughly is true.

(Authier 1992:332)

If we accept that topicalization targets the same position as negative inversion, then the position of topicalization is [Spec, CP], contra Baltin's IP-adjunction approach.

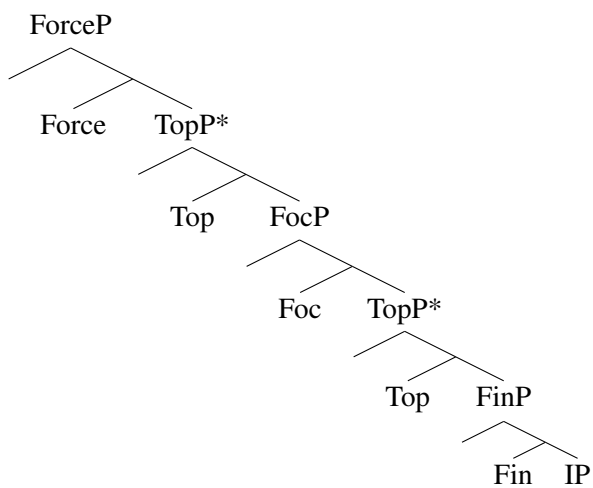
2.2.3 The Split-CP approach

Introduction of the split CP framework by Rizzi (1997) provided further support for the idea that topicalization is a CP phenomenon. Rizzi regards CP as the interface area between the sentence (syntax) and discourse. It is here where various discourse-related notions including topic, focus, and speaker/hearer information are reflected onto the structure of the sentence. In this regard, the traditional single-CP layer is

insufficient to encode such complex information. Neither is the homogeneous double CP-layer suggested in the CP-recursion approach (Authier 1992, Watanabe 1993). This inadequacy calls for a richer and fine-grained structure at the CP area with a strict hierarchy among functional heads: the split CP layer.

This new framework brought about a wealth of research around various phenomena of the syntax-pragmatics interface, with many researchers aiming to uncover the exact "cartography" of some area of the split CP. Topicalization is one of the phenomena that were taken into consideration in the original inception of the framework in Rizzi (1997), where two recursive topic layers (TopP*) are seated below and above focus (FocP).

(27)



(Rizzi 1997:297)

Since then, researchers of this cartographic research program have aimed to fine-tune the position of topic based on word order data from various languages. Their main assumption is that there is a universal hierarchy of functional projections, although languages parametrically differ as to which heads are projected and how they are implemented in the morphosyntax and prosody (Rizzi and Bocci 2017).

Haegeman (2004, 2006) surveys data from a relatively wide range of languages (English, Gungbe, Korean, Japanese, and a number of romance languages including Italian, French, Catalan and Spanish). She allocates the higher TopP for English topicalization and the lower TopP for Romance clitic left dislocation. Her argument is based on the asymmetry between two types of subordinate clauses, namely *peripheral adverbial clauses* (PACs) and *central adverbial clauses* (CACs). PACs modify the speech act and are more loosely related to the main clause. CACs, on the other hand, modify the event denoted by the main clause and are more "tightly" embedded into it. Below is her example of a Korean PAC (28a) and a CAC (28b).[4]

(28) a. ku chayk-ul cohaha-n-ta-myen way kukes-ul sa-ci anh-ni?
that book-ACC like-PRES-DEC-if why that-ACC buy-NMZ not.do-Q
"If you like that book, why don't you buy it?"

b. (ku-ka) i chayk-ul ilk-umyen/ilk-ess-umyen ku-nun ama
(he-NOM) this book-ACC read-if/read-PST-if he-TOP probably

[4]Glosses are presented as they are in the original source. Some choices of terminology do not coincide with mine and thus differ from other parts of the thesis – e.g. *ta* as a declarative (DEC) morpheme instead of a sentence final particle (SFP).

ku yenghwa-lul poko siphe hal kes-i-ta.
that movie-ACC see want will-DEC

"If he reads/read this book, he will probably want to see that movie."

(Haegeman 2006:1656)

The antecedent of the conditional clause (28a), *if you like that book*, is a PAC. It is understood as a premise that should be considered in relation to the whole question speech act of the main clause. On the contrary, the antecedent of (28b) is a CAC which denotes an event that is likely to become the cause of the consequent.

The distinction between CACs and PACs may seem to be mainly semantic, but in fact the two also differ in their syntactic behavior. One of Haegeman's main observations is that argument topicalization is only available in PACs.

(29) Topicalization inside CACs

a. * If these exams you don't pass you won't get the degree.

b. * While her book Mary was writing this time last year, her children were staying with her mother.

(30) Topicalization inside PACs

a. His face not many admired, while his character still fewer felt they could praise.

b. If aphids we did not worry about, snails we did.

(Haegeman 2004:159-160)

Haegeman takes a split-CP approach in explaining the difference between the two types of adverbial clauses: CACs are structurally reduced, only projecting up to FinP, while PACs project up to ForceP.[5] Since PACs include additional functional projections, there is "room" within a PAC to license topicalization that CACs lack. The fact that expressions of epistemic modality, which are anchored to the speaker, are also exclusively available in PACs provides additional evidence to her claim.

(31) a. ?? PAC: John works best while his children are *probably/might* be asleep.

b. * CAC: Mary accepted the invitation without hesitation after John *may* have accepted it.

There have been some attempts to apply the split-CP approach to topicalization in East Asian languages as well. Lee (2008) argues that the order of topic and focus in Korean as C-Topic-Focus-T, although she takes Chomsky's (1977) original view that topicalization is base-generation of topic at TP (Topic Phrase) plus movement of a *wh*-operator. Saito (2012) examines the structural hierarchy among Japanese com-

[5]While Haegeman does embrace Rizzi's general framework, she postulates a different ordering of functional heads which is suggested in Roussou (2000).

(i) [Sub [Topic/Focus [Force [Fin]]]]

Here topic and focus are above Force, so the projection of the latter does not structurally *entail* projection of the former. However, she does argue on conceptual grounds that topicalization strongly *depends* on the presence of Force. Referring to Bayer (2001), she describes topicalization as a type of speech act, which involves Force or Speaker Deixis in Haegeman's terms. Thus she strongly argues that "English ... does not have any alternative way of relating a fronted topic to the associated clause" and that "[T]hat topicalization is not available in central adverbial clauses in English is a consequence of the absence of the projection of the head Force." (Haegeman 2004:170) The takeaway is that in both her and Rizzi's original analysis the presence of Force guarantees the availability of Topic, although on different grounds.

plementizers *no*, *ka*, and *to*, establishing the hierarchy of (32). He locates Japanese topic, which is accompanied by the suffix *wa*, above *no* and below *ka*.

(32) [... [... [... [... [$_{TP}$ ...] Finite (*no*)] (Topic*)] Force (*ka*)] ... Report (*to*)]

While it is beyond the scope of this study to determine the exact cartographic structure regarding the position of TopP, it should be noted that some influential approaches of the split-CP camp imply that the presence of Force either conceptually or structurally guarantees the availability of topic (Rizzi 1997, Haegeman 2004, 2006, Saito 2012).

2.3 The syntax of factive clauses and embedded topicalization

Since Kiparsky and Kiparsky's (1970) influential paper (henceforth K&K), it has been acknowledged that a certain group of predicates which select for clausal complements displays a distinct semantic behavior. Here are some examples of the verbs that they identified as peculiar.

(33) *regret, be aware (of), comprehend, ignore, forget, make clear*

In order for utterance of these predicates and their clausal complement clause to be felicitous, the truth of the complement must be presupposed by the speaker. K&K's original example effectively illustrates this point (wording is slightly altered for ease of explanation).

(34) a. Factive: It is significant that John has been found guilty.

b. Non-factive: It is likely that John has been found guilty.

As for (34b), it is imaginable that some speaker utters (34b) without believing in the truth of the complement clause. In other words, it is possible for the speaker to utter (34b) without believing that John has been found guilty. On the contrary, a reasonable speaker would not utter (34a) without believing that John has been found guilty. This is because the semantics of (34a) requires the truth of the complement clause to be *presupposed* by the speaker. One of the tests for diagnosing factivity is placing a contradictory sentence or clause right after the sentence of interest (Karttunen 1971, Basse 2008, Kastner 2015 among others). As demonstrated below, the contradicted clause leads to infelicity only when the contradicted proposition is introduced by a factive predicate (35b).

(35) a. I thought that the building collapsed, but it didn't.

b. # I regretted that the building collapsed, but it didn't.

(Kastner 2015:159)

K&K emphasize that presupposition of the truth of some clause is different from *asserting* the truth of that clause (See also Hooper and Thompson (1973), Stalnaker (1975) and Hegarty (1992)). They compare the two examples below, where the truth of the complement clause in (36a) is asserted while in (36b) it is presupposed.

(36) a. It is true that John is ill.

b. It is odd that the door is closed.

The speaker directly expresses belief in the truth of the complement clause *by means of* uttering (36a) – that is, the speaker *asserts* its truth. On the contrary, when uttering (36b), the speaker does believe that the door is closed but this is not what is being foregrounded. Here the speaker asserts an opinion about the content of the embedded clause: namely that it is odd. The factive clause *that the door is closed* is said to "lack assertion."

The semantic peculiarity of factive clauses is also reflected in their syntactic particularity. One such example is that they do not allow topicalization inside them.

(37) a. John said [that Sue, Bill doesn't like.]

b. * John regretted [that *Gone with the wind*, we went to see.]

(Watanabe 1993)

Hooper and Thompson (1973) link the unavailability of topicalization inside factive clauses with their lack of assertion by pointing out that since topicalization is a speaker-oriented phenomenon that involves some point-of-view or assertion originating from the speaker.

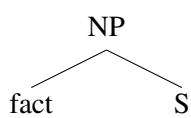
Researchers have theorized that factive clauses have a different structure from non-factive clauses, employing different grammatical mechanisms to account for their syntactic behavior. The rest of this section is structured as follows. Subsec-

tion 2.3.1 introduces the classic account of K&K on the structure of factive clauses. Subsections 2.3.2 and 2.3.3 reviews two main approaches to the structure of factive clauses: the operator approach and the clause-size approach. It is also explained how the lack of topicalization inside factive clauses (37) is accounted for under the two approaches. Subsection 2.3.4 presents Korean examples of factive clauses and examines the prediction of the two approaches on Korean data.

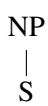
2.3.1 The original account of K&K

K&K's initial analysis is that factive clauses are complements of the noun *fact*. Non-factive embedded clauses, on the other hand, are simply S.

(38) Factive



(39) Non-factive



Only factive clauses can be a complement of an overt noun *fact*. This is, they argue, supporting evidence of their analysis. Consider (40-41).

(40) a. I regret the fact that John is ill.

b. I want to make clear the fact that I don't intend to participate.

c. * I assert the fact that I don't intend to participate.

(41) a. I regret the fact of John's being ill.

b. You have to keep in mind the fact of his having proposed several alternatives.

c. * We may conclude the fact of his having proposed several alternatives.

(Kiparsky and Kiparsky 1970: 345, 356)

The deep structure of (38) is left bare in (40a-b) and (41a-b). If the a-sentences in (40-41) undergo *fact*-deletion, they will still be grammatical sentences.

(42) a. I regret that John is ill.

b. I regret John's being ill.

K&K's analysis is the most straightforward way to structurally account for the semantics of factive clauses. However, it is difficult to reconcile this with recent theoretical advances under the minimalist program where every movement should be strictly motivated. There are no clear motivations for the transformations that apply liberally to derive the factive structure. Furthermore, if factive clauses are structurally embedded under an NP as the complement of the noun *fact*, they are expected to display complex NP island effects just like actual complex NPs. This, however, is not always the case as English object extraction is marginally available in factive clauses but not in actual complex NP islands. In the examples below, *wh*-extraction of the argument *what* is marginally available from true factive clauses (43) but not from complement clauses of NPs (44).

(43) a. ? What do you regret that John stole?

b. * Why do you remember that John stole the cookies? (Basse 2008)

(44) a. * What do you remember the fact that John stole?

b. * Why do you remember the fact that John stole the cookies?[6]

(Kastner 2015)

Such problems have led researchers to search for a syntactic explanation that better suits the data and the theoretical mechanisms while retaining the intuition of K&K.

2.3.2 The operator approach

Instead of placing an actual noun *fact*, some accounts postulate a semantically motivated operator at the edge of factive clauses. Melvold (1991) assumes that factivity is a definite description of an event, and that a "definiteness" operator is present at [Spec, CP]. It is this operator that is also responsible for the weak islandness of factive clauses demonstrated in (43).

Watanabe (1993) adopts and develops Melvold's idea, combining it with the double-CP layer insight of Authier (1992) introduced in section 2.2.2. He assumes following Authier that topic is indeed a CP phenomenon as affirmed in section 2.2, and that the higher [Spec, CP] is reserved for *wh*-movement. Then, the operator approach would predict a conflict between the factive operator and the topicalized ele-

[6] As Kastner clarifies, it is ungrammatical in the "low" reading as intended. This of course also applies to Basse's example (43b).

ment since it targets the same position where the operator is located: the lower of the two [Spec, CP]s as in (45a). Hence in sentences such as (45b) arguments like *Mary* cannot be topicalized, since it would have to land at the spot where the definiteness operator is already situated.

(45) a. ... [$_{CP}$ [$_{C'}$ that [$_{CP}$ **Topic/Op** [$_{C'}$ ∅ [$_{Agr\text{-}sP}$]]]]]

b. John regrets [$_{CP}$ that [$_{CP}$ Op [$_{Agr\text{-}SP}$ he fired Mary]]]

A more recent implementation of the operator approach is presented by Haegeman and Ürögdi (2010). They also posit a similar operator at the edge of factive clauses. Their "event" operator is compared to a *wh*-operator raising in interrogative clauses and relative clauses: both contain a [+Q] feature and raise from within the clause to the edge. Instead of appealing to competition for the same position, their account relies on featural relativized minimality (RM). Featural RM stems from the insight of Rizzi (2001, 2004) who states that the "identical structural type" requirement of the original RM (Rizzi 1990) should be redefined. In the original framework, structural type refers one of three subtypes: head, A specifier, and A-bar specifier (Rizzi 1990: 6-7). Some justification for the A/Abar distinction in defining structural type is provided from French (Obenauer 1983). Consider first that the French *wh*-specifier *combien* can either be pied-piped with the rest of the DP or be extracted alone.

(46) a. [Combien de livres]$_i$ a-t-il consultés t_i? (French)

(*lit.*) "How many of books did he consult?"

b. Combien$_i$ a-t-il consulté [t_i de livres]?

(*lit.*) "How many books did he consult of books?"

(Obenauer 1983, cited from Rizzi 1990)

The trace of the DP *combien de livres* in (46a) is theta-governed[7] by the verb *consulté(s)*, while the trace of the extracted specifier *combien* is antecedent-governed by the specifier itself. In either case, the trace is adequately governed and there are no locality constraint violations. However, the grammaticality judgment changes with addition of the VP-initial quantifying adverb *beaucoup*, as demonstrated in (47).

(47) a. [Combien de livres]$_i$ a-t-il beaucoup consultés t_i?

(*lit.*) "How many books did he a lot consult?"

b. * Combien$_i$ a-t-il beaucoup consulté [t_i de livres]?

(*lit.*) "How many did he a lot consult of books?"

(Obenauer 1983, cited from Rizzi 1990)

[7]RM shares basic assumptions with the Empty Category Principle or ECP (Chomsky 1981, Stowell 1981) in that it limits movement by means of government requirements on nonpronominal empty categories such as traces.

(i) ECP: A nonpronominal empty category must be
   a. Theta-governed, or
   b. antecedent-governed. (Stowell 1981, cited from Rizzi 1990:4)

RM deviates from the original versions of ECP in that potential interveners for antecedent government are restricted to elements of the same structural type.

Rizzi explains that the position of *beaucoup* qualifies as an A-bar landing position. If we adopt the views of RM and view the adverb as an intervener for antecedent government of *combien*, the ungrammaticality of (47b) can be explained in a clear way. Hence Obernauer's example has been considered as empirical evidence supporting the argument that RM is configured to rely on the A/A-bar distinction.

However, it has been later observed that a locality constraint such as that of (47b) does not hold for all A-bar chains. Consider (48).

(48) a. [Combien de livres]$_i$ a-t-il attentivement consultés t_i?

(*lit.*) "How many books did he carefully consult?"

b. Combien$_i$ a-t-il attentivement consulté [t_i de livres]?

(*lit.*) "How many did he carefully consult of books?"

(Obenauer 1994, cited from Rizzi 2004)

Despite being in the exact same position as *beaucoup*, the adverb *attentivement* here does not lead to ungrammaticality. Apparently, just placing any phrase at the designated A-bar position does not lead to intervention for A-bar chains.

The asymmetry shown in (47-48) suggests that the configuration of RM should go beyond a definition that relies on the simpler distinction of A/A-bar positions and opt for a more fine-tuned definition regarding structural types. One of the suggested solutions to this problem is to look into the featural makeup of the antecedents and interveners. As Starke (2001) points out, we can now subdivide structural identity

into three cases (Starke 2001: 7-8):

(a) The antecedent and the intervener are of the exact same featural makeup:

$*\alpha_i \quad ... \quad \alpha_j \quad ... \quad \alpha_i$

(b) The intervener is featurally *less* specified than the antecedent:

$\alpha\beta \quad ... \quad \alpha \quad ... \quad \alpha\beta$

(c) The intervener is featurally *more* specified than the antecedent:

$*\alpha \quad ... \quad \alpha\beta \quad ... \quad \alpha$

The α elements consist of the typical quantificational features such as *wh, neg, foc, quantificational adverb* that make up weak islands. As for $\alpha\beta$s, their feature set is a subclass of α's: there is something "extra" in their feature set as compared to that of αs. Of the three cases above, (a) and (c) lead to ungrammaticality but not (b). This is because the more specific $\alpha\beta$ chain can be formulated via two potential movements: α-movement or $\alpha\beta$-movement. α can be an intervener for α-movement but not for $\alpha\beta$-movement. On the contrary, $\alpha\beta$ in (c) is by definition an intervener for α-movement.

Researchers have made various suggestions as to which interpretative aspect this extra specification represented by β corresponds to. One of the prominent theories is related to referentiality, presupposition, and *d*(iscourse)-linking (Cinque 1990, Rizzi 2004). Starke (2001) provides examples with contextual backgrounds, which we will

not investigate in detail here, demonstrating how extraction out of weak islands such as *whether*-islands are only possible when there is an existential presupposition of the referent of the extracted phrase.[8] In more informal terms, "there exist[s] some entity which the interlocutor has in mind as a referent." (Starke 2001:13)

Returning to the operator approach of Haegeman and Ürögdi (2010), the authors reunite the above insight of Starke with the notion of topic. As introduced in the beginning of section 2.1, topic is a relationally given discourse element. It has been directly or indirectly mentioned in the previous discourse and is now considered old information. Its existence is inevitably presupposed by the discourse participants, and this qualifies topic as the "specified" $\alpha\beta$ element described above. (The authors follow Boeckx and Jeong (2004) and call the specifying discoursal feature δ – see (50).) This entails that topic is a potential intervener for non *d*-linked, normal quantifier movement (α-movement in the above formalization).

One can now analyze the status of the event operator under Starke's framework as well. The operator quantifies over the whole event, is responsible of the weak island status of factive clauses, and cannot move over topic, which is an $\alpha\beta$ element (Here formalized by the feature Q). If we consider it an α element in Starke's terms, the aforementioned phenomena of factive clauses can be accounted for in an elegant manner. The operator starts from above TP and raises to the clause typing position in

[8] He formalizes such extractions as $Q\beta$-movement as opposed to Q-movement; this is obviously reminiscent of the α-$\alpha\beta$ contrast above.

the CP area. (FP is an assumed functional projection initially housing the operator.)

(49) $[_{CP}\ Op_{Q} \ldots [_{FP}\ t\ [_{TP}\ V \ldots\ ]]]$

However, if an XP is topicalized and fronted as in (50) below, the event operator raising past the topic is blocked by intervention.

(50) * $[_{CP}\ Op_{Q}\quad XP_{\delta+Q} \ldots [_{FP}\ t\ [_{TP}\ V \ldots]]]$

The topic is featurally richer than the operator, so the latter raising over the former causes a locality violation. Since the presence of the event operator at [Spec, CP] is crucial for factive clauses, the derivation crashes. To summarize, under the operator approach to factive clauses, topicalization is unavailable because it collides with the crucial factive operator which must be present at the clause edge.

2.3.3 The clause-size approach

A different line of research focuses on a more structural difference between factive and non-factive clauses. This study concentrates on Haegeman's earlier works (2004, 2006) as an example of a structural approach to factive clauses. It has already been introduced in section 2.2 that she adopts the split-CP framework of Rizzi (1997), differentiating CACs from PACs. Under her view, factive clauses are CACs and are structurally reduced. Non-factive clauses correspond to PACs, projecting more structure than factive clauses. To support this correspondence, she demonstrates that the

syntactic difference between CACs and PACs also holds between factive and non-factive clauses. For example, speaker-oriented adverbs, which are exclusively available in PACs, are not available in factive clauses.

(51) * John regrets that Mary probably/obviously/unfortunately did not attend the meeting. (Haegeman 2006:1664)

Once we accept the CAC-PAC distinction for factive and non-factive clauses, the lack of topicalization inside factive clauses is easily explained. It has already been demonstrated that within English conditional clauses, topicalization is only available in PACs (29-30, repeated here as 52-53)

(52) Topicalization inside CACs

a. * If these exams you don't pass you won't get the degree.

b. * While her book Mary was writing this time last year, her children were staying with her mother.

(53) Topicalization inside PACs

a. His face not many admired, while his character still fewer felt they could praise.

b. If aphids we did not worry about, snails we did.

(Haegeman 2004:159-160)

This is because PACs project up to ForceP; projection of ForceP guarantees licensing

of Topicalization at TopicP (see footnote 5 in section 2.2.3). If factive clauses correspond to CACs such as (52) and non-factive clauses correspond to PACs like (53), they are expected to behave similarly in terms of topicalization: it is only available inside non-factive clauses.

Note that in assuming additional structure for non-factive clauses here, she is going against K&K's intuition that factive clauses are the marked case and non-factives the unmarked ones. Non-factive clauses project "extra structure" that anchors the uttered proposition to the speaker: the functional head *Speaker Deixis* just below the subordinating head *Sub*. This account is strongly reminiscent of the early observations by Hooper and Thompson (1973) that factive clauses "lack assertion" – here they literally lack structure that is related to assertion, unlike non-factive clauses which do project assertion-related structure. The insight that non-factives are what require additional structure is also shared by De Cuba (2007) and Nichols (2001), among others.

2.3.4 Factive clauses in Korean

In Korean, factive clauses and non-factive clauses are noticeably different in structure. They are selected by different complementizers: *ko* for non-factive, *kes* for factive.

(54) a. Non-factive verb *mitta* "believe"

Hani-ka [Minho-ka cwungkwuke-lul calha-n-ta ko]
Hani-NOM Minho-NOM Chinese-ACC do.well-PRES-SFP Comp
mit-nun-ta.
believe-PRES-SFP

"Hani believes that Minho is fluent in Chinese."

b. Factive verb *nollawehata* "be surprised"

Hani-ka [Minho-ka cwungkwuke-lul calha-nun kes]-ul
Hani-NOM Minho-NOM Chinese-ACC do.well-ADN Comp-ACC
nollawehay-ss-ta.
be.surprised-PST-SFP

"Hani was surprised that Minho is fluent in Chinese."

Unlike the non-factive complementizer *ko*, the status of *kes* has been much debated. Its pronominal status is agreed upon by most researchers; however, whether it also functions as a complementizer is controversial. Since *kes* is freely interchangeable with actual NP/DPs, a considerable number of researchers have classified it as a pro-form even in complementizer-like environments. However, this study follows Kang (2006), who provides some evidence that *kes* is a true complementizer when it appears in factive clauses and cleft constructions. It displays different syntactic behaviors from nominals that appear in the same position.

First, *kes* in a cleft clause cannot co-occur with the plural marker *tul*. The pro-form *kes* can, of course, be pluralized.

(55) Mary-ga manna-n *kes*-(*tul)/saram-(tul)-un John-kwa Sally-i-tta
Mary-NOM meet-PRE *kes*/person-TOP John-and Sally-be-DECL
"It is John and Sally that Mary met."

Kang considers *kes* in cleft sentences to be complementizers. The fact that the noun *saram* "person" but not *kes* agrees with the plural noun *John-kwa Sally* "John and Sally" supports this view.

Another piece of evidence comes from optional postposition dropping in cleft constructions. As for clefts with *kes*, postpositions attached to the focalized element can optionally be dropped. This does not hold for true nominals such as *cangso*.

(56) a. Mary-ga John-ul manna-n *kes*-un i-kongwon-(eseo)-i-tta.
Mary-NOM John-ACC meet-PRE *kes*-TOP this-park-at-be-DECL

b. Mary-ga John-ul manna-n *cangso*-nun i-kongwon-(*eseo)-i-tta.
Mary-NOM John-ACC meet-PRE place-TOP this-park-at-be-DECL
"It is at this park that Mary met John."

Kang explains that the reason why the postposition *eseo* is unavailable in (56b) is because *cangso*, an NP, cannot be co-referential with *i-kongwan-eseo*, a PP. The fact that *kes* in (56a) behaves differently from the nominal *cangso* is indirect evidence that it is a type of complementizer, not a simple pronominal element.

Returning to the factive clause of Korean, the crucial observation is that there are two different forms of factive clauses. Observe the pair in (57). (57b) contains additional functional elements at the clause edge that (57a) does not: the tense marker *n* and the sentence final particle (SFP) *ta*. While there have been different viewpoints on the exact status of *ta*, here it is called SFP, a theory-neutral term employed by Pak (2008). She regards *ta* to be a sentential force marker along with *nya* for in-

terrogatives. Under the clause-size approach or split-CP framework, (57b) would be considered "larger" in size than (57a).

(57) a. Hani-ka [Minho-ka cwungkwuke-lul calha-nun kes]-ul
Hani-NOM Minho-NOM Chinese-ACC do.well-ADN Comp-ACC
nollawehay-ss-ta.
be.surprised-PST-SFP

b. Hani-ka [Minho-ka cwungkwuke-lul calha**-n-ta**-nun
Hani-NOM Minho-NOM Chinese-ACC do.well-PRES-SFP-ADN
kes]-ul nollawehay-ss-ta.
Comp-ACC be.surprised-PST-SFP
"Hani was surprised that Minho speaks Chinese well."

Both (57a) and (57b) are factive clauses since the factive presupposition is required in both cases for their utterances to be felicitous. To demonstrate this point, I adopt the diagnostic test introduced in example (35) where I place a semantically contradicting sentence right after the sentence of interest. The contradiction does not lead to infelicity after a non-factive sentence (58a) while it does for both forms of factive sentences (58b-c). The semantic anomaly of both (58b-c) suggest that they are equally factive.

(58) a. Hani-ka [Minho-ka cwungkwuke-lul calha-n-ta-ko]
Hani-NOM Minho-NOM Chinese-ACC do.well-PRES-SFP-Comp
malhay-ss-ta.
say-PST-SFP

Haciman Minho-nun sasil cwungkwuke-lul mosha-n-ta.
but Minho-TOP actually Chinese-ACC do.bad-PRES-SFP

"Hani said that Minho speaks Chinese well. But actually, Minho doesn't

speak Chinese."

b. Hani-ka [Minho-ka cwungkwuke-lul calha-nun kes]-ul
Hani-NOM Minho-NOM Chinese-ACC do.well-ADN Comp-ACC
nollawehay-ss-ta.
be.surprised-PST-SFP

Haciman Minho-nun sasil cwungkwuke-lul mosha-n-ta.
but Minho-TOP actually Chinese-ACC do.bad-PRES-SFP

c. Hani-ka [Minho-ka cwungkwuke-lul calha-n-ta-nun
Hani-NOM Minho-NOM Chinese-ACC do.well-PRES-SFP-ADN
kes]-ul nollawehay-ss-ta.
Comp-ACC be.surprised-PST-SFP

Haciman Minho-nun sasil cwungkwuke-lul mosha-n-ta.
but Minho-TOP actually Chinese-ACC do.bad-PRES-SFP

"Hani was surprised that Minho speaks Chinese well. #But actually, Minho doesn't speak Chinese."

From a semantic perspective, it has been suggested that the two forms of factive clauses differ in the source of the factivity (Lee 2017). The factive presupposition of (57b) is motivated internally, or from the speaker's own cognitive sources, while that of (57a) is motivated from indirect sources such as hearsay or reported information. (See section 5.2 for a more thorough discussion.)

In terms of syntax, however, there have not been much discussion about this duality. Previous cross-linguistic literature on factivity have not entertained the possibility of a two-way system of factive clauses. In the context of previous literature reviewed in this chapter regarding topicalization and factivity, there is a question that ensues:

how does topicalization work in each of the cases in (57)?

Sections 2.3.2 – 2.3.3 have reviewed two main approaches on the structure of factive clauses: the operator approach and the clause-size approach. For most languages with only one type of factive structure, the operator approach and the clause size approach provide similar predictions: topic is not licensed inside factive clauses. However, matters are different for Korean since neither approach has considered the possibility of two different structures of factive clauses as shown in (57). Following the logic of the two approaches, it is expected that they provide different predictions for topic licensing inside Korean factive clauses, especially for those that include an SFP. Consider the pair of sentences below. These are identical to (57) except that the embedded subject of the factive clause is topicalized: (59b) projects more structure and is "larger" than (59a) in split-CP terms.

(59) a. Hani-ka [Minho-**nun** cwungkwuke-lul calha-nun kes]-ul
Hani-NOM Minho-TOP Chinese-ACC do.well-ADN Comp-ACC
nollawehay-ss-ta.
be.surprised-PST-SFP

b. Hani-ka [Minho-**nun** cwungkwuke-lul calha-**n-ta**-nun
Hani-NOM Minho-TOP Chinese-ACC do.well-PRES-SFP-ADN
kes]-ul nollawehay-ss-ta.
Comp-ACC be.surprised-PST-SFP
"Hani was surprised that Minho (topicalized) speaks Chinese well."

Under the operator approach, both (59a) and (59b) contain a definiteness or event operator at the clause edge. Following Watanabe's (1993) approach, the topicalized sub-

ject *Minho-nun* ungrammatically occupies the [Spec, CP] where the operator should lie. In the featural RM account of Haegeman and Ürögdi (2010), the factive operator cannot raise beyond *Minho-nun* in both (59a) and (59b), again resulting in ungrammaticality.

The clause-size approach predicts differently. The tense marker and SFP of (59b) are functional heads at the periphery of the clause; they would each correspond to an independent projection of their own. While the cartography of the Korean right periphery is far from identified, the SFP *ta* can arguably be located at ForceP since it encodes sentential force (Portner 2004, Pak 2008). Then the range of projections would differ for (59a) and (59b). Since there is no compelling evidence for extra structure at the edge of the embedded clause in (59a), it can be understood to project up to FinP – identical to what Haegeman has suggested for factive clauses. *Minho-nun* would cause ungrammaticality because the factive clause does not include TopP to license it. On the other hand, with the inclusion of an SFP, the factive clause in (59b) would project up to ForceP. Projection of ForceP entails projection of TopP. This means that (59b), despite being factive, is able to license *Minho-nun*.

To summarize, the operator approach expects both (59a) and (59b) to be equally ungrammatical while the clause-size approach predicts that only the former would be ungrammatical. Since the two approaches present different predictions for Korean factive clauses, the next logical step would be to empirically verify whether

the predictions of the two approaches are borne out. Since judgments on syntactico-pragmatic phenomena such as factivity and topicalization are not very clear-cut, an acceptability judgment experiment was carried out. A well-controlled formal experiment yields collective, reliable and robust data on subtle judgment tasks. One can also look at grammaticality in a more gradient viewpoint, which allows to formulate a more precise theory on subtle data. The specific design and result of the experiment is presented in section 3.

3 The Main Experiment

3.1 Logic and design of the experiment

The main experiment aimed to verify whether the presence of the SFP *ta* ameliorates topic licensing inside factive clauses. This research question can be implemented by a 2×2 factorial design: an intersection of two categorical factors with two levels each. The relevant factors are SFP and TOPIC, each containing the levels {*With SFP* – *Without SFP*} and {*Topic-marked* – *Nominative-marked*} respectively. A combination of one level from each factor constitutes an experimental *condition*; there would be four conditions in a 2×2 factorial design. Below is a table visualizing the 2×2 design and the corresponding sentences for each condition. The columns correspond to the two levels of the SFP factor while the rows correspond to the two levels of the TOPIC factor.

Table 1: 2×2 factorial design for the main experiment

	With SFP	Without SFP
Nominative-marked	(60a)	(60b)
Topic-marked	(60c)	(60d)

(60) a. (= 57b)

Hani-ka [Minho-ka cwungkwuke-lul calha-**n**-**ta**-nun
Hani-NOM Minho-NOM Chinese-ACC do.good-PRES-SFP-ADN
kes]-ul nollaweha-yss-ta.
Comp-ACC be.surprised-PST-SFP

b. (= 57a)

Hani-ka [Minho-ka cwungkwuke-lul calha-nun kes]-ul
Hani-NOM Minho-NOM Chinese-ACC do.good-ADN Comp-ACC
nollaweha-yss-ta.
be.surprised-PST-SFP

c. (= 59b)

Hani-ka [Minho-**nun** cwungkwuke-lul calha-**n**-**ta**-nun
Hani-NOM Minho-TOP Chinese-ACC do.good-PRES-SFP-ADN
kes]-ul nollaweha-yss-ta.
Comp-ACC be.surprised-PST-SFP

d. (= 59a)

Hani-ka [Minho-**nun** cwungkwuke-lul calha-nun kes]-ul
Hani-NOM Minho-TOP Chinese-ACC do.good-ADN Comp-ACC
nollaweha-yss-ta.
be.surprised-PST-SFP

"Hani was surprised that Minho speaks Chinese well."

The nominative-marked sentences, (60a-b), are expected to be grammatical under both the operator approach and the clause-size approach since there is no topicalization to degrade grammaticality. The operator approach predicts (60c-d) to be both significantly worse than (60a-b): a *significant main effect of* TOPIC in experimental terms. The clause-size approach would predict a significant difference between the grammaticality of (60c-d) but not between (60a-b). In other words, it would predict a *significant interaction of* SFP:TOPIC – the presence/absence of SFP significantly

affects grammaticality, but only when the embedded subject is topic-marked.

The experiment was an acceptability judgment task: participants read each stimulus as presented on screen and judged its grammaticality on a 7-point Likert scale (from 1 to 7, with 1 labeled "very unnatural" and 7 labeled "very natural" in Korean).

In designing the items, it was the semantics of the verbs that were most strictly controlled since this is what determines the factivity of the complement clause. The matrix verbs were limited to verbs that were deemed factive in previous literature (Kiparsky and Kiparsky 1970, Hooper and Thompson 1973, Hegarty 1992; Lee, Nam, and Kang 1998 on Korean). Additionally, there have been observations since Karttunen (1971) that there are two classes of factive verbs, namely emotive factive verbs and cognitive factive verbs. Taking this into consideration, four of the eight token sets were designed with an emotive verb while the rest included cognitive factive verbs. The specific verbs are presented below. (See Appendix B for the full list of stimuli.)

- cognitive factive verbs

  kiekhaynayta (remember), *nwunchichayta* (notice), *alanayta* (find out), *kkaytatta* (realize)

- emotive factive verbs

  kippehata (be glad), *sulphehata* (be sad), *nollawehata* (be surprised), *anthakkawehata* (regret)

As for the embedded clauses, their semantic contents were chosen so that they would easily license sentential topic. This is an effort to exclude other possible confounds that affect topic licensing. Predicates of the embedded clauses were limited to individual-level predicates. Individual-level predicates portray characteristic or generic information about the subject and are readily compatible with topic, which encodes given information (Chierchia 1995, Lee 2003a). If topic licensing is limited in non-SFP environments in spite of such topic-friendly semantics, this limitation can be attributed to purely syntactic constraints that this study aims to uncover. Below is the list of the eight embedded predicates used in the stimuli.

- *mwunhak-ul silhehata* (dislike literature)
- *kay-lul mwusewehata* (be afraid of dogs)
- *misin-ul mitta* (believe in superstitions)
- *kohyang-ul kuliwehata* (miss one's hometown)
- *wuntong-ul culkita* (enjoy sports)
- *kacok-ul salanghata* (love one's family)
- *cwungkwuke-lul calhata* (be fluent in Chinese)
- *chinkwutul-ul cilthwuhata* (be jealous of one's friends)

It has been explained in section 2.1.1 that the Korean topic marker *nun* is ambiguous between a contrastive topic marker and non-contrastive topic marker. The study only aimed to investigate non-contrastive sentential topic, so measures were taken to suppress the contrastive reading. The stimuli were presented in the form of question-answer pairs where the question provides adequate context so as to induce the non-contrastive reading for the *nun*-marked DP in the answer. This is inspired from the way Korean non-contrastive topic is introduced in Jun (2015). An example is given in (61).

(61) Q: Minho-ka ecey nwukwu-lul manna-ss-ni?
Minho-NOM yesterday who-ACC meet-PST-INT
"Who did Minho meet yesterday?"

A: Minho-nun ecey Tongho-lul manna-ss-ta.
Minho-TOP yesterday Tongho-ACC meet-PST-DECL
"As for Minho, he met Tongho yesterday." (Jun 2015:183)

Minho-nun in (61B) qualifies as sentential topic defined in section 2.1.1. Since it is mentioned in the immediately preceding discourse (61A), the speaker of (61B) must be cognitively aware of it (cf. example (5)); hence it is referentially given. It is relationally given in contrast with *Tongho-lul*, which is the focus and relationally new information of the sentence. Furthermore, *Minho-nun* serves as the entry under which new information about *Tongho-lul* is to be updated. Since *Minho-ka* and *Minho-nun* in (61) are co-referential and not in a part-whole relation, a contextually-induced alternative set for *Minho-nun* is unavailable or at least much less salient than the

"money-coin" pair in (6). Since the alternative set is what evokes a contrastive reading, the lack of such a set results in the suppression of the contrastive interpretation of *nun*. For these reasons, it was judged that question-answer pairs such as (61) are an effective means to induce the non-contrastive sentential topic reading of *nun*-marked DPs in Korean. Experimental stimuli were designed accordingly – see (64-67) and Appendix B for concrete examples.

Another confound that affects judgment is speakers' avoidance of consecutive uses of the same morphological case marker. In (62), the Korean nominative case marker *ka* is attached to two adjoining DPs.

(62) Nay-**ka** Hani-**ka** kacok-ul sarangha-nun kes-ul kkaydal-ass-e.
I-NOM Hani-NOM family-ACC love-ADN COMP-ACC realize-PST-SFP
"I realized that Hani loves her family."

In a preparatory pilot study, the target stimuli were designed like (62). Several participants of the pilot study gave feedback after the experiment that the consecutive appearance of two *ka*s stood out as ungrammatical and led them to give the sentence a low score. While the psychological mechanism behind this avoidance of repetition is not identified, the phenomenon itself is well known as the Obligatory Contour Principle (OCP).[9] The concept originated from phonology (Leben 1973, McCarthy 1986 among others) but has been extended to other modules of the linguistic process including syntax under the name *generalized OCP* (Mohanan 1994).

[9] See Hiraiwa (2010) for an overview.

(63) The Generalized OCP: Universal

Identical elements (melodic units/formatives) are disallowed in adjacent units.

(not absolute) (Mohanan 1994:212, cited from Hiraiwa 2010)

In order to prevent syntactic OCP effects in the stimuli, a temporal adverb (either *onul* "today", *ecey* "yesterday" or *akka* "a while ago") was inserted between the two subjects.

Below is a sample token set with a corresponding question-answer pair for each condition. (The question is identical across all conditions.)

(64) Nominative-marked + With SFP

Q: Ney-ka onul Hani-ey tayhayse mwues-ul kkaytal-ass-e?
you-NOM today Hani-LOC about what-ACC realize-PST-SFP

"What did you realize about Hani today?"

A: Nay-ka onul Hani-ka kacok-ul sarangha**-n-ta**-nun kes-ul
I-NOM today Hani-NOM family-ACC love-PRES-SFP-ADN COMP-ACC
kkaydal-ass-e.
realize-PST-SFP

"I realized today that Hani loves (her) family."

(65) Nominative-marked + Without SFP

Q: Ney-ka onul Hani-ey tayhayse mwues-ul kkaytal-ass-e?
you-NOM today Hani-LOC about what-ACC realize-PST-SFP

"What did you realize about Hani today?"

A: Nay-ka onul Hani-ka kacok-ul sarangha-nun kes-ul
I-NOM today Hani-NOM family-ACC love-ADN COMP-ACC
kkaydal-ass-e.
realize-PST-SFP

"I realized today that Hani loves (her) family."

(66) Topic-marked + With SFP

Q: Ney-ka onul Hani-ey tayhayse mwues-ul kkaytal-ass-e?
you-NOM today Hani-LOC about what-ACC realize-PST-SFP

"What did you realize about Hani today?"

A: Nay-ka onul Hani-nun kacok-ul sarangha**-n-ta**-nun kes-ul
I-NOM today Hani-TOP family-ACC love-PRES-SFP-ADN COMP-ACC
kkaydal-ass-e.
realize-PST-SFP

"I realized today that Hani loves (her) family."

(67) Topic-marked + Without SFP

Q: Ney-ka onul Hani-ey tayhayse mwues-ul kkaytal-ass-e?
you-NOM today Hani-LOC about what-ACC realize-PST-SFP

"What did you realize about Hani today?"

A: Nay-ka onul Hani-nun kacok-ul sarangha-nun kes-ul
I-NOM today Hani-TOP family-ACC love-ADN COMP-ACC
kkaydal-ass-e.
realize-PST-SFP

"I realized today that Hani loves (her) family."

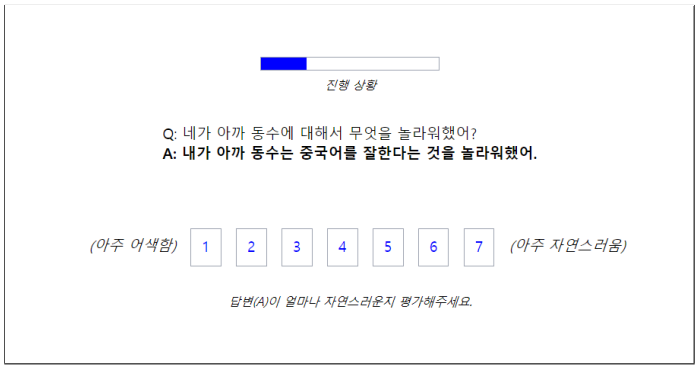
Eight token sets like the one presented in (64-67) were designed as experimental stimuli. They were put on a Latin square design so that each participant was exposed

to only one stimulus from each token set, or two stimuli per condition. Participants also saw 24 filler items of varying grammaticality. They were asked to judge seven announced practice items so that they could become familiar with the tasks. Judgments on practice items were not included in the statistical analyses. For each of the four lists formed by the Latin square, four different pseudorandomized orderings were produced to offset any ordering effects. This resulted in a total of sixteen possible survey orderings.

61 participants were recruited online. Their first language is Korean, but the study did not collect information on whether they had any knowledge of languages other than Korean. Their ages range from 20 to 53. Participants were able to withdraw at any time during the experiment without any disadvantage; this fact was clearly stated on the introduction screen. After completion of the experiment, participants were paid for their time and effort.

The experiments were conducted online on *Ibex Farm*: participants were assigned to one of the sixteen survey orderings and were provided with the appropriate *Ibex Farm* URL to access their assigned experiment. The following figure is a sample screenshot of the experiment platform that participants saw.

Figure 1: Sample screenshot of the main experiment on *Ibex Farm*



3.2 Results

Raw score averages and standard deviations of each of the four conditions are presented below in table 2.

Table 2: Mean and standard deviation of the four conditions of the main experiment

	With SFP		Without SFP	
	Mean	SD	Mean	SD
Nominative-marked	5.58	1.43	5.25	1.55
Topic-marked	4.88	1.84	3.95	1.84

Raw scores of the stimuli and fillers were z-score transformed for each participant prior to analysis. This is a standardization measure to correct ceiling/floor effects or skews each individual participant may have in terms of score giving (Sprouse 2011, Sprouse et al. 2011, 2016).

Using the *lmer* function of the *lmerTest* package in R, a linear mixed effects

model was constructed. The fixed effects were SFP, TOPIC, and the interaction term SFP:TOPIC. The random effects were ITEM nested in TOKENSET (`tokenSet/item`) and PARTICIPANT. As table 3 demonstrates, there are both a significant main effect of TOPIC ($p < 0.01$) and a significant interaction of SFP:TOPIC ($p < 0.05$).

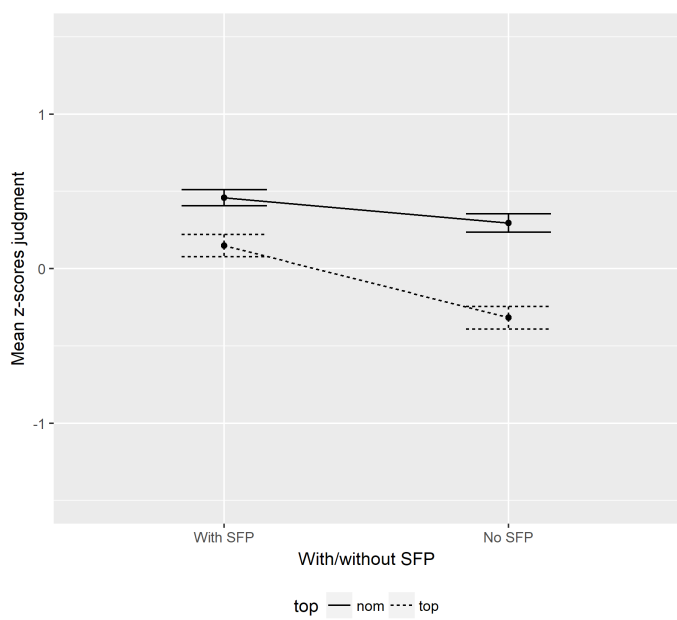
Table 3: Significance of the fixed effects of the main experiment

	Estimate	Std. Error	p-value
Intercept	0.44564	0.13429	0.00786**
SFP	−0.14426	0.08563	0.10787
Topic	−0.28485	0.08564	0.00343**
SFP:Topic	−0.33811	0.12112	0.01139*

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Figure 2 shows an interaction plot of the factors SFP and TOPIC.

Figure 2: Interaction plot for the main experiment



A post-hoc comparison test was run using the *emmeans* function of the *emmeans* package in R. This test runs a t-test for every possible pair of conditions and checks whether there is a significant difference between their means. Out of the six possible t-tests, the four relevant tests are presented below.

Table 4: Post-hoc comparison test of the main experiment

Contrast	p-value
[Nominative + With SFP] – [Nominative + Without SFP]	0.6464
[Topic + With SFP] – [Topic + Without SFP]	0.0001**
[Nominative + With SFP] – [Topic + With SFP]	0.0203*
[Nominative + Without SFP] – [Topic + Without SFP]	< 0.0001***

- p-value adjustment: Bonferroni method for 6 tests (Confidence level used: 0.95)
- Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

The first two rows show whether a change in the SFP level made a significant difference when the TOPIC level is controlled. Differences are observed only within *Topic-marked* stimuli ($p < 0.001$), and not for *Nominative-marked* ones ($p = 0.6464$). This is expected, since there is a significant interaction of SFP:TOPIC but no significant main effect of SFP. (See table 3)

The bottom two rows represent differences arising from changing the TOPIC level when the SFP level is controlled. In this case, both *With SFP* and *Without SFP* levels show a significant difference. Again, this can be explained by the fact that there is a significant main effect of TOPIC as well as a significant interaction.

The results of the main experiment demonstrate both significant main effect of

TOPIC and significant interaction of SFP:TOPIC. Hence, the experiment provides empirical support for both the operator approach and the clause-size approach. The factivity of the main clause predicate affects embedded topic licensing to a certain level, but more so in smaller-sized embedded clauses. There is one caveat in accepting the operator approach, though. The overall z-score of the [Topic + With SFP] condition is higher than zero (see figure 2). This means that participants rated sentences corresponding to this condition higher than their individual averages, although they rated them significantly lower than sentences corresponding to the [Nominative + With SFP] condition. It may be rather misguiding to bluntly label the [Topic + With SFP] condition as "ungrammatical" as the operator approach predicts.

4 The Follow-up experiment

4.1 Logic and design of the experiment

The follow-up experiment extends the logic of the main experiment to island constructions. It has been observed that, like factive clauses, islands in general disallow topic licensing inside themselves. (It should be noted that the island effects mentioned here indicate ungrammaticality of the licensee *within* the island, not cases where elements inside the island cannot *scramble out* to the main clause.)

(68) a. * The fact that Bill, Mary likes makes John very jealous.

(Authier 1992)

b. * When her regular column she began to write for The Times, I thought she would be OK.

(Haegeman 2006)

c. * John-ga kono hon-wa yonda hito-ni atta. (Japanese)
John-NOM this book-TOP read man-to met
"John met the man who this book, read."

(Maki et al. 1999)

This may not always be the case in Korean, where some islands mandatorily include an SFP (*ta*, *nya*) while others do not.

(69) a. Complex NP Island (With SFP)

Nay-ka [Nahuy-ka misin-ul mit-nun-ta-nun
I-NOM Nahuy-NOM superstition-ACC believe-PRES-SFP-ADN
somwun]-ul cenha-yss-e.
rumor-ACC deliver-PST-SFP.

"I delivered the rumor that Nahuy is superstitious."

b. *wh*-island (With SFP)

Nay-ka [Cinsoli-ka mwues-ul mwuseweha-nya-ko]
I-NOM Cinsoli-NOM what-ACC be.afraid-SFP-Comp
cilmwunha-yss-e.
question-PST-SFP.

"I asked what Cinsol is afraid of."

c. Adjunct (*because*) island (Without SFP)

Nay-ka [Socini$_k$-ka sengsilhaki ttaymwuney] kyay$_k$-lul
I-NOM [Socin-NOM diligent because] her$_k$-ACC
chingchanha-yss-e.
praise-PST-SFP.

"I praised Socin because she is diligent."

d. Relative clause island (Without SFP)

Nay-ka [Hyeyswu-ka sanun t_k] tongney$_k$-lul
I-NOM [Hyeswu-NOM live t_k] neighborhood$_k$-ACC
mwusiha-yss-e.
look.down-PST-SFP

"I looked down on the neighborhood where Hyeswu lives."

Results from the main experiment showed a significant main effect of TOPIC and a significant interaction of SFP:TOPIC). Since the operator approach (Melvold 1991,

Watanabe 1993, Haegeman and Ürögdi 2010) is specific to factive clauses, it may not be directly applicable to island constructions. The clause-size approach, on the other hand, is expected to hold for islands as well because the split-CP framework covers various types of clausal constructions. If this prediction is on the right track, islands with SFP (69a-b) are expected to show significantly better grammaticality than those without (69c-d) when the embedded subject is topicalized. Or, in experimental terms, a significant interaction of SFP:TOPIC is expected to be observed.

Just as for the main experiment, the follow-up experiment was designed as an acceptability judgment task involving a 7-point Likert scale. The target stimuli were built based on the same 2×2 factors SFP and TOPIC. However, since the level of the SFP factor (*With SFP, Without SFP*) is dependent on the specific token set (*i.e.* the island type), it is impossible to construct a token set representing all four possible conditions. Hence, the Latin square design is unavailable for this experiment. Therefore it was decided that the experiment be designed as a full repeated measure - all participants were exposed to all of the 32 target stimuli.

Other formal and semantic considerations remained identical to the main experiment. Stimuli were again provided in a question-answer format, where the question provides adequate context to facilitate sentential topic licensing and to suppress the contrastive reading of the *nun*-marked subject. The target stimuli included a temporal adverb between the matrix subject and the embedded subject. Embedded clause

predicates were individual-level predicates to ensure semantic topic licensing. Below are the sample stimuli for each island type. Notice the alternation of the nominative case marker *ka* and the topic marker *nun* in the answers. (Refer to Appendix C for the full list of stimuli.)

(70) **Complex NP island** (island *With SFP*)

Q: Ney-ka onul Nahuy kwanlyenhayse iyakiha-yss-e? Mwusun
you-NOM today Nahuy about talk.do-PST-SFP which
iyaki-lul ha-yss-e?
talk-ACC do-PST-SFP

"Did you talk about Nahuy today? What did you say?"

A: Nay-ka onul Nahuy-{ka/nun} misin-ul
I-NOM today Nahuy-{NOM/TOP} superstition-ACC
mit-nun-ta-nun somwun-ul cenha-yss-e.
believe-PRES-SFP-ADN rumor-ACC deliver-PST-SFP

"I delivered the rumor today that Nahuy is superstitious."

(71) ***wh*-island** (island *With SFP*)

Q: Ney-ka onul Somi kwanlyenhayse cilmwunha-yss-e? Mwusun
you-NOM today Somi about question.do-PST-SFP which
cilmwun-ul ha-yss-e?
question-ACC do-PST-SFP

"Did you ask a question about Somi today? What did you ask?"

A: Nay-ka onul Somi-{ka/nun} nwukwu-lul salanghanu-nya-ko
I-NOM today Somi-{NOM/TOP} who-ACC love-SFP-COMP
mwulepo-ass-e.
ask-PST-SFP

"I asked today who Somi loves."

(72) **Relative clause island** (island *Without SFP*)

Q: Ney-ka akka Caywu kwanlyenhayse iyakiha-yss-e? Mwusun
you-NOM while.ago Caywu about talk.do-PST-SFP which
iyaki-lul ha-yss-e?
talk-ACC do-PST-SFP

"Did you talk about Caywu a while ago? What did you say?"

A: Nay-ka akka Caywu-{ka/nun} tani-nun hoysa-lul
I-NOM while.ago Caywu-{NOM/TOP} work.for-ADN company-ACC
pinanha-yss-e.
criticize-PST-SFP

"I criticized a while ago the company that Caywu works for."

(73) **Adjunct (*because*) island** (island *Without SFP*)

Q: Ney-ka ecey Socini kwanlyenhayse iyakiha-yss-e? Mwusun
you-NOM yesterday Socin about talk.do-PST-SFP which
iyaki-lul ha-yss-e?
talk-ACC do-PST-SFP

"Did you talk about Socin yesterday? What did you say?"

A: Nay-ka onul Socini-{ka/nun} sengsilha-ki ttaymwuney kyay-lul
I-NOM today Socin-{NOM/TOP} diligent-*ki* because her-ACC
chingchanha-yss-e.
praise-PST-SFP

"I praised Socin yesterday because she is diligent."

70 participants were recruited online.[10] Again, their first language is Korean, but the study did not collect information on whether they had any knowledge of languages other than Korean. The age range is also identical: 20-53. Participants were able to withdraw at any time during the experiment without any disadvantage; this

fact was clearly stated on the introduction screen. After the experiment, participants were paid for their time and effort.

Four token sets (lexicalizations) were designed for each island type. Each token set consisted of two versions: one with the case-marked subject and the other with the topic-marked subject. This resulted in eight stimuli for one island type, a total of 32 stimuli. As clarified above in the Item Design section, every participant was exposed to the same set of items. Participants also saw 48 filler items of varying grammaticality, and seven announced practice items. Four different pseudorandomized orderings were produced to offset any ordering effects. Again, the experiments were conducted online on *Ibex Farm*: participants were provided the corresponding *Ibex Farm* URL to access the experiment of their assigned ordering.

4.2 Results

In the follow-up experiment, a total of four island types were tested: two islands with an SFP (CNP and *wh* islands) and two islands without an SFP (relative clause and

[10]Due to difficulties in participant recruiting, 42 participants, or 60 percent of the sample, participated in both the main experiment and the follow-up experiment. A Welch two-sample t-test was conducted to check whether the overlapping and non-overlapping population showed different z-score means. For the entire sample (42 participants versus 28 participants, judgment for fillers as well as stimuli), the test did not support a significant difference between the two groups. ($t = 0.13569$, $p = 0.89$)

The same test was conducted for each condition; only a marginal difference was suggested between the two samples for the condition [Topic + With SFP].

	Nom. W/O SFP	Nom. With SFP	Top. W/O SFP	Top. With SFP
p-values	0.31	0.52	0.78	0.05

Based on the above results, it was concluded that the two samples can be considered as homogeneous and may be analyzed as such.

adjunct islands). Table 5 summarizes the mean and standard deviation of the raw scores for each type of island, with the subject either nominative-marked or topic-marked.

Table 5: Mean and std. dev. of the eight conditions (SFP × ISLANDTYPE) of the follow-up experiment

	CNP		*wh*		RC		Adjunct	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Nominative-marked	5.46	1.49	6.19	1.21	5.86	1.35	4.33	1.68
Topic-marked	5.47	1.43	6.24	1.13	2.00	1.33	4.03	1.72

The four island types are collapsed into two levels of the SFP factor. Complex NP island and *wh*-island are collapsed into the *With SFP* level while Relative clause island and Adjunct (*because*) island are collapsed into the *Without SFP* level. Table 6 presents the corresponding mean and standard deviation of raw scores.

Table 6: Mean and std. dev. of the four conditions (SFP×TOPIC) of the follow-up experiment

	With SFP (CNP + *wh*)		Without SFP (RC + Adjunct)	
	Mean	SD	Mean	SD
Nominative-marked	5.83	1.40	5.09	1.70
Topic-marked	5.86	1.34	3.01	1.84

Two different statistical analyses were carried out for this experiment. The first analysis is equal to that of the main experiment: verification of any main effects and significant interaction of the two factors SFP and TOPIC. The second analysis is a complement of the first, whereby each island type is compared separately.

4.2.1 2×2 factorial analysis

Again, raw scores were z-score transformed prior to linear regression modeling. A linear mixed effects model was constructed using the *lmerTest* package in R to verify whether there is a significant interaction of SFP:TOPIC. Hence, the two factors SFP, TOPIC and their interaction were included as fixed factors. Random effects included ITEM nested inside TOKENSET (`tokenSet/item`) and PARTICIPANT as in the main experiment, but this time TOKENSET nested in ISLANDTYPE (`islandType/tokenSet`) was also added as a random effect. The results yielded a significant interaction of SFP:TOPIC ($p < 0.001$) without any significant main effects.

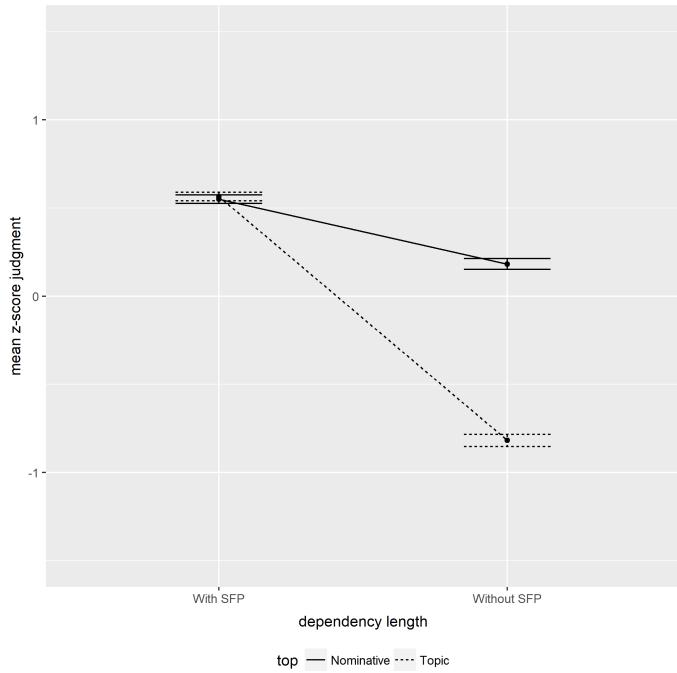
Table 7: Significance of the fixed effects of the follow-up experiment

	Estimate	Std. Error	p-value
Intercept	0.5506	0.1660	0.030*
SFP	−0.3683	0.2343	0.1918
Topic	0.0141	0.1796	0.9380
SFP:Topic	−1.0150	0.2540	0.0005***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Figure 3 is an interaction plot of the fixed factors.

Figure 3: Interaction plot of the fixed effects of the follow-up experiment



As with the main experiment, a post-hoc comparison test was executed for the follow-up experiment as well. A change in SFP level exclusively affected the *Topic-marked* level, and a change in TOPIC level exclusively affected *Without SFP* level. This is an expected outcome given the presence of a significant interaction and the absence of significant main effects.

Table 8: Post-hoc comparison test of the follow-up experiment

Contrast	p-value
[Nominative + With SFP] – [Nominative + Without SFP]	1.0000
[Topic + With SFP] – [Topic + Without SFP]	0.0265*
[Nominative + With SFP] – [Topic + With SFP]	1.0000
[Nominative + Without SFP] – [Topic + Without SFP]	0.0004***

- p-value adjustment: Bonferroni method for 6 tests (Confidence level used: 0.95)
- Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

4.2.2 By-island analysis

The by-island analysis is a planned comparison test on each island type. It retains the individual island types unlike the 2×2 factorial analysis, which collapses island types into levels of the SFP factor. The test verifies whether there is a significant difference between nominative-marked and topic-marked stimuli within each island type.

The comparison was conducted on a linear mixed effects model, with ISLAND-TYPE and TOPIC as fixed effects, and ITEM nested inside TOKENSET and PARTICIPANT as random effects. The planned comparison test was conducted in the same manner as the post-hoc tests above, using the *emmeans* package in R.

Table 9: By-island planned comparison of the follow-up experiment

Contrast (Nominative-marked vs. Topic-marked)	p-value
Complex NP island	1.0000
wh island	1.0000
Relative clause island	< 0.0001***
Adjunct (*because*) island	0.4422

- p-value adjustment: Bonferroni method for 28 tests (Confidence level used: 0.95)
- Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

There was no significant difference between nominative-marked and topic-marked stimuli for CNP islands nor for *wh*-islands. This is not unexpected since the post-hoc test for aggregated *With SFP* islands did not yield significant results. (See table 8)

What is surprising is the difference between relative clause islands and adjunct (*because*) islands. The relative clause islands behave as expected, showing a significant difference between nominative-marked and topic-marked stimuli. This is not the case for the adjunct islands. It can be inferred from these results that the significant difference between nominative-marked and topic-marked stimuli for *Without SFP* islands in the 2×2 factorial analysis were due to the significant difference within relative clause islands.

The overall results of the follow-up experiment provide some empirical support for the clause-size approach. The grammaticality of topicalized subjects in islands with SFP (CNP, *wh*) and their ungrammaticality in relative clause islands is expected. As for Adjunct (*because*) islands, more investigation is needed. It may be that the

Korean morpheme *ki* is situated in a high functional projection (e.g. SFP) whose presence entails the availability of topic licensing. An additional experiment that examines other forms of Korean *because*-clauses (e.g. *se* or *nikka*) instead of *ki* may lead to fruitful discussion. (See section 5.4 for relevant discussion.)

5 Discussion

Two different approaches on the structural analysis of factive clauses were presented: the operator approach and the clause-size approach. The results in general support the clause-size approach, which expects factive clauses and islands to be able to license topic when an SFP is present. However, there are remaining issues regarding the specific structures of factive clauses and islands in Korean. This section briefly covers four discussion points that suggest the need for a deeper investigation, answering potential questions that may be raised.

5.1 The structural projection of "small" factive clauses

The structural projection of "large" factive clauses has, to a degree, become clear. Assuming the clause-size approach and the underlying split-CP framework, this study suggests on conceptual and empirical grounds that the large factive clauses must project up to ForceP. But it is relatively unclear how much the smaller factive clauses project. The argument that the split-CP framework is appropriate for Korean data lies on the assumption that even "smaller" factive clauses do project a CP element, namely FinP. If the small factive clauses in Korean are shown not to project FinP but in fact stop lower, for example below TP, then the argument for a split-CP framework approach is weakened.

However, there is indirect evidence that the "small" factive clauses are still larger than TP. In the experimental stimuli, it was impossible to include an overt tense mor-

pheme inside the factive clauses because the embedded predicates were limited to individual-level predicates. These predicates do not go together with temporal adverbs or tense markers due to their semantics. (Chierchia 1995) When the semantics of the verb is altered to be stage-level, tense and aspect can overtly be expressed as in (74). Furthermore, (74) displays an additional projection, *te*, outside the tense morpheme *(e)ss*; this suggests that the subordinate clause here is larger than TP.

(74) Hani-ka [Minho-ka/*nun chayk-ul sse-ss-te-n kes]-ul
Hani-NOM Minho-NOM/TOP book-ACC write-PST-*te*-ADN *kes*-ACC
alanay-ss-ta.
discover-PST-SFP
"Hani discovered that Minho had written a book."

Assuming that *(e)ss* here is a tense morpheme denoting past tense, it is likely that *te* is situated at a functional projection higher than T. One possible candidate is the evidential projection, which is observed to be outside of tense and aspect expressions such as *(e)ss* and *keyss* (Kwon 2012a,b for analysis of *te* as a retrospective evidential marker in Korean).

Similar observations hold in relative clause islands, which demonstrated striking ungrammaticality in the follow-up experiment when its internal subject was topicalized.

(75) Hani-ka [[Minho-ka/*nun sse-ss-te-n] chayk]-ul ilk-ess-ta.
Hani-NOM Minho-NOM/TOP write-PST-*te*-ADN book-ACC read-PST-SFP
"Hani read the book that Minho had written."

Again, the relative clause projects the evidential *te* in addition to what was presented as the experimental stimuli. The presence of *te* outside of *(e)ss* excludes the possibility where the small factive clauses and relative clauses are simply TPs; the evidential projection is a C element that comes higher than past tense in Cinque's (1999) hierarchy. Yet the topicalized embedded subject *Minho-nun* cannot be licensed inside of them. These facts call for a split-CP framework, which explains the availability of topicalization by the size of the structural projection at the C area, to be applied to the data.

5.2 Interpretation of "large" factive clauses

Previous literature has almost exclusively discussed small factive clauses. The large ones projecting SFP are, to the best of my knowledge, not recognized outside of the context of Korean linguistics. Even among literature on Korean, the issue has not yet been much discussed, especially in syntax. From a semantic perspective, Lee (1976, 2017) defines the small factive clauses as "internal" and the large ones as "external".

> The external type thus is used to imply that the embedded clause proposition p has been asserted elsewhere in the context via a discourse move and that it has been conveyed in an indirect evidence acquisition to the attitude holder as something like public or shared knowledge to be in the Common Ground. (Lee 2017: 6)

The internal type grounds its truth from the speaker's internal motives (feelings, experience, etc.) while the external type presupposes the truth based on "objectified" or indirect sources. This reflects the insight that the SFP + adnominal *ta-nun* is grammaticalized from the quotative *tako hanun* (Chang 1987, Lee 2017).

One possible explanation is to reconcile this semantic insight with the assumed function of ForceP: to encode assertion or speaker deixis (Haegeman 2006). If the large factive clause is external, meaning that its factive presupposition is grounded on indirect or reportative sources, it can also be said that the content of the factive clause has been *asserted* elsewhere by some other speaker who has provided the source for the presupposition. Hence the inclusion of SFP and consequent topic licensing despite the factivity of the embedded clause.

5.3 Presupposition triggers and the source of factivity

A fundamental question that arises from the duality of Korean factive clauses is where the "source" or trigger of factivity lies within the factive clause. Since Levinson (1983:181-184) had laid out the concept of *presupposition triggers*, or expressions that evoke a certain presupposition when used in an utterance, researchers have investigated the presupposition trigger for factive presuppositions (Djärv et al. 2016 for a recent example). The factive presupposition trigger indicated by Levinson (1983) is the predicate. Indeed, the lexical semantics of verbs such as *regret* or *notice* require that their complement clauses hold true. However, at least for Korean, there is another

element that is involved in triggering the factive presupposition: the complementizer *kes*.

Kang (2000) explores the various cases where the semantics of complementizers interacts with the meaning of the complement clause in Korean. The complementizer *ko* always appears with non-factive propositions while *kes* can appear with both factive and non-factive propositions. Hence it may seem that the distribution of *kes* subsumes that of *ko*, but this is not always the case. Consider (76-77). The a-sentences are grammatical with the complementizer *ko* but the b-sentences with *kes* are more awkward.

(76) a. Cengswu-nun [Myenghuy-ka ttena-ss-ta-ko]
Cengswu-TOP Myenghuy-NOM leave-PST-SFP-Comp
sayngkakha-yss-ta.
think-PST-SFP

b. ?? Cengswu-nun [Myenghuy-ka ttena-ss-ta-nun kes]-ul
Cengswu-TOP Myenghuy-NOM leave-PST-SFP Comp-ACC
sayngkakha-yss-ta.
think-PST-SFP
"Cengswu thought that Myenghuy left."

(77) a. Cengswu-nun [Myenghuy-ka ttokttokha-ta-ko]
Cengswu-TOP Myenghuy-NOM be.smart-SFP-Comp
mit-nun-ta.
believe-PST-SFP

b. ?? Cengswu-nun [Myenghuy-ka ttokttokha-ta-nun kes]-ul
Cengswu-TOP Myenghuy-NOM be.smart-SFP Comp-ACC
sayngkakha-yss-ta.
believe-PST-SFP

"Cengswu thought that Myenghuy left."

(Kang 2000:215)

Cases like (76-77) contribute in delineating the exact property of *kes*, which is a difficult task since *kes* can appear in a wide range of environments. Kang explains that *kes* does not fit in the above sentences because their matrix predicates are psyche verbs that express subjective thinking of the agent *Cengswu*. The implied subjectivity of the compliment clause does not accord with the meaning of *kes*. He claims that when a clause is headed by *kes*, it is understood as an object of abstract observation or a matter of discussion. In agreement with Lee et al. (1998), he defines the semantics of *kes* as *thematization*. [11]

(78) provides further evidence that *kes* is a thematizing complementizer. In the sentences below, the matrix predicate is *mitta* "believe" just like in (77). However, the use of *kes* in (78) does not lead to semantic anomaly because the compliment denotes a proposition that has already been much discussed by others: the existence of God.

(78) Cengswu-nun [sin-i concayha-n-ta-nun kes]-ul
Cengswu-TOP God-NOM exist-PRES-SFP-ADN Comp-ACC
mit-nun-ta / mit-ci anh-nun-ta
believe-PRES-SFP / believe-NMZ do.not-PRES-SFP
"Cengswu believes/does not believe that God exists."

[11] More accurately, it is what he calls *ta-nun kes* clauses (*kes* clauses with an SFP) that are thematized. Under his view, *kes* clauses without an SFP undergo type coercion when they are subordinated under factive predicates so that they too can refer to a proposition.

(Kang 2000:215)

The concept of thematization is in a similar vein with some of the qualities of factive clauses. They are characterized by their lack assertion, which implies the lack of viewpoint or speaker anchoring. Thematization is taking an objective stance towards the propositional content, which also implies a lack of subjective viewpoint. If the meaning of *kes* does indeed involve thematization, there is a possibility that it too is a factive presupposition trigger in the sense of Levinson (1983). While it is beyond the scope of this paper to provide a full-fledged formal analysis for the factive interpretation of the two types of Korean factive clauses, the multiplex nature of factive presupposition triggers in Korean may be accountable for the complicated aspects of factivity in the language.

5.4 The clause size of adjunct *because* islands

It was verified in the results of the follow-up experiment that the grammaticality of Korean adjunct *because* islands does not change significantly when the subject is topicalized. In fact, the mean and standard deviation barely change. A related issue is the multitude of structures in Korean corresponding to the *because* clause. "Reason" can be introduced by the connectors *ki ttaymwuney*, *se*, or *nikka*. *Nikka* is excluded in this discussion because it anchors to a level higher than the other two: it is only available in epistemic judgments and cannot easily be embedded nor take scope under

conditionals (Park 2015).

Experimental stimuli of the follow-up experiment were constructed only with *ki ttaymwuney* clauses. Raw mean scores for both nominative-marked or topic-marked conditions are formed around 4.3 – almost right in the middle. The scores suggest that topicalization is marginally available in *ki ttaymwuney* clauses. An explanation consistent with the clause-size approach would be to consider *ki ttaymwuney* as larger than it initially seems: possibly equivalent to PACs or large factive clauses, but the fact that SFPs cannot be included inside it casts doubt on the idea.

(79) * Hani-ka [Minho$_i$-ka chincelha-ta-ki ttaymwuney kyay]$_i$-lul
Hani-NOM Minho-NOM kind-SFP-*ki* because him-ACC
chingchanha-yss-ta.
praise-PST-SFP
"Because Minho$_i$ is kind, Hani praised him$_i$."

Comparison of *ki ttaymwuney* with *se* may prompt fruitful discussion. They display different syntactic behaviors in several aspects; for example, the former but not the latter allows tense morphemes inside of them.

(80) a. Hani-ka [Minho-ka tachie-ss-ki ttaymwuney] kyay$_i$-lul
Hani-NOM Minho$_i$-NOM be.injured-PST-*ki* because him$_i$-ACC
tolpwa-ss-ta.
take.care-PST-SFP

b. * Hani-ka [Minho-ka tachie-ss-se] kyay$_i$-lul
Hani-NOM Minho$_i$-NOM be.injured-PST-*se* him$_i$-ACC
tolpwa-ss-ta.
take.care-PST-SFP
"Because Minho$_i$ was injured, Hani took care of him$_i$."

As pointed out in section 5.1, the presence of tense and aspect is insufficient to guarantee topic licensing. However, the asymmetry of (80a) and (80b) hints at a size difference between the two types of *because* islands. An informal survey of seven native speakers of Korean asked for their judgment on the same experimental stimuli but with *se* instead of *ki ttaymwuney*. Raw score averages are 4.21 for items with nominative-marked embedded subjects and 3.71 for items with topic-marked embedded subjects – not a small difference. Future research could include a controlled experiment with *se* adjunct clauses and report if there are any differences in tendency between *se* and *ki ttaymwuney*. This may contribute to clarifying the position of *se* and *ki ttaymwuney* and their respective functional projections, and furthermore to formulating a theory on the interaction of topicalization and island clause size.

6 Conclusion

The present study has compared two theoretical approaches or models on factive clauses: the operator approach and the clause-size approach. The objective was to empirically verify the two models against data that can tease them apart. Topicalization was chosen as the criterion for comparison, since it has been much discussed in terms of its availability in embedded environments including factive clauses. Most of the languages already documented in related literature only demonstrate a single form of factive clause. Hence it is difficult to distinguish the predictions of the two approaches or to identify a better model. Korean, on the other hand, possesses two different types of factive clauses which are of different sizes. They minimally differ in the presence or absence of a handful of functional particles. The semantically motivated operator approach does not expect a difference in topic licensing for the two structures. The syntactically defined clause-size approach, on the other hand, expects the larger factive clauses to better license topicalization than their smaller counterparts.

An acceptability judgment experiment of 2×2 factorial design was conducted in order to compare the actual grammaticality of the two types of clauses. In terms of the experimental factors, the operator expects a significant main effect of TOPIC while the clause-size approach expects a significant interaction of SFP:TOPIC. Results of the main experiment coincides with the prediction of both approaches, suggesting

the need for a hybrid or more complex theory of embedded topicalization in factive clauses.

Based on the findings from the main experiment, a follow-up experiment was designed to verify whether the clause-size approach can also cover island data from Korean. Islands have traditionally been known to disallow topicalization inside of themselves. (Authier 1992, Maki et al. 1999). However, Korean islands differ in whether they include an SFP at the clause edge. If the results reveal that islands with SFPs license topicalization significantly better than those without SFPs, they can serve as additional evidence that the clause-size approach is on the right track. The results provide partial support for this view. Topicalization inside complex NP islands and *wh*-islands, which include an SFP inside them, were significantly more grammatical than topicalization inside relative clause islands which do not project an SFP. However, the difference in grammaticality was not as salient for adjunct (*because*) islands. This unexpected observation calls for additional research on structure of *because* clauses in Korean.

The contribution of this study in a wider context is its empirical support for the split-CP approach of Rizzi (1997). Korean is characteristic in its abundant use of particles at the clause edge to express various and complex pragmatic information including evidentiality, honorifics, and modality. Furthermore, these particles line up at the clause edge under a strict hierarchy: an arbitrary change in their order results

in strong ungrammaticality. The results of the two experiments imply that the clause-edge particles are indeed important syntactic elements. Their absence or presence have significant consequences that can affect the grammaticality of the whole clause. Despite such adequacy of the split-CP approach, there have not been abundant research in this direction. Compared to some influential studies in Japanese and Chinese (Endo 2007, Saito 2012, Tsai 2015), similar attempts in Korean have been relatively rare. Future research may consider exploring the cartographic approach to provide concrete explanations on various phenomena of the syntax-pragmatics interface in Korean.

References

Authier, Jean-Marc. 1992. Iterated CPs and Embedded Topicalisation. *Linguistic Inquiry* 23:329–336.

Badan, Linda, and Francesca Del Gobbo. 2010. On the syntax of topic and focus in chinese. In *Mapping the left periphery*, ed. Paola Benincà and Nicola Munaro. Oxford University Press.

Baltin, Mark R. 1982. A landing site theory of movement rules. *Linguistic Inquiry* 13:1–38.

Basse, Galen. 2008. Factive Complements as Defective Phases. *Proceedings of WCCFL 27* 54–62.

Bayer, Josef. 2001. Asymmetry in emphatic topicalization. *Audiatur vox sapientiae* 52:15–47.

Boeckx, Cedric, and Youngmi Jeong. 2004. The fine structure of intervention in syntax. In *Issues in current linguistic theory: A festschrift for hong bae lee*, 83–116. Kyungchin Seoul.

Büring, Daniel. 2003. On D-Trees, Beans, and B-Accents. *Linguistics and Philosophy* 26:511–545.

Chafe, Wallace. 1976. Givenness, contrastiveness, definiteness, subjects, topics and point of view. In *Subject and topic*, ed. Charles N. Li, 27–55. Academic Press.

Chang, Kyunghee. 1987. Kwukeuy wanhyengpoceluy haysek (The interpretation of Korean non-defective complement clauses). *Kwukehak (Journal of Korean Linguistics)* 16:487–519. [Written in Korean].

Chierchia, Gennaro. 1995. Individual-level predicates as inherent generics. In *The generic book*, 176–223. University of Chicago Press.

Chomsky, Noam. 1977. On Wh-Movement. In *Formal syntax*, ed. Peter Culicover, Thomas Wasaw, and Adrian Akmajian, 71–132. London: Academic Press.

Chomsky, Noam. 1981. *Lectures on government and binding: The pisa lectures*. Walter de Gruyter.

Chomsky, Noam. 1986. *Barriers*. The MIT Press.

Cinque, Guglielmo. 1990. *Types of $\bar{A}$-dependencies*. MIT press.

Cinque, Guglielmo. 1999. *Adverbs and Functional Heads: A Cross-Linguistic Perspective*. Oxford University Press.

De Cuba, Carlos Francisco. 2007. On (non)factivity, clausal complementation and the CP-field. Doctoral dissertation, Stony Brook University.

Djärv, Kajsa, Jérémy Zehr, and Florian Schawrz. 2016. Cognitive vs. emotive factives: An experimental differentiation. *Proceedings of Sinn und Bedeutung 21* .

Emonds, Joseph. 1969. Root and Structure-preserving Transformations. Doctoral Dissertation, MIT.

Emonds, Joseph. 1976. *A transformational approach to English syntax: Root, structure-preserving, and local transformations*. Academic Press Inc.

Endo, Yoshio. 2007. *Locality and Information Structure: A cartographic approach to Japanese*. John Benjamins.

Frascarelli, Mara, and Roland Hinterhölzl. 2007. Types of topics in German and Italian. In *On information structure, meaning and form*, ed. Kerstin Schwabe and Susanne Winkler, 87–116. John Benjamins Publishing Company.

Gundel, Jeanette. 1985. Shared knowledge and topicality. *Journal of Pragmatics* 9:83–107.

Gundel, Jeanette. 1988. Universals of topic-comment structure. In *Studies in syntactic typology*, ed. Michael Hammond, Edith A. Moravcsik, and Jessica Wirth, 209–239. Amsterdam: John Benjamins Publishing Company.

Gundel, Jeanette, and Thorstein Fretheim. 2004. Topic and focus. In *The handbook of pragmatics*, ed. Laurence Horn and Gregory Ward, 175–196. Oxford: Blackwell.

Haegeman, Liliane. 2004. Topicalization, CLLD and the Left Periphery. *ZAS Papers in Linguistics* 1:157–192.

Haegeman, Liliane. 2006. Conditionals, factives and the left periphery. *Lingua* 116:1651–1669.

Haegeman, Liliane, and Barbara Ürögdi. 2010. Referential CPs and DPs: An operator movement account. *Theoretical Linguistics* 36.

Halliday, Michael. 1967. Notes on transitivity and theme in English: Parts 1 and 2. *Journal of Linguistics* 3:37–81, 199–244.

Halliday, Michael, and Christian Matthiessen. 2004. *An introduction to functional grammar*. Routledge, third edition.

Hegarty, Michael. 1992. Adjunct Extraction and Chain Configurations. Doctoral dissertation, MIT.

Higgins, Francis. 1973. On J. Emonds' Analysis of Extraposition. *Syntax and Semantics* 2.

Hiraiwa, Ken. 2010. The syntactic OCP. *Tokyo conference on psycholinguistics* 4:35–56.

Hooper, Joan, and Sandra Thompson. 1973. On the Appilcability of Root Transformations. *Linguistic Inquiry* 4:465–497.

Huang, James C.-T., Audrey Y.-H. Li, and Yafei Li. 2009. *The Syntax of Chinese*. Cambridge University Press.

Jun, Youngchul. 2006. Korean Marker 'NUN' for Contrastive Topic and Contrastive Focus. *HAN-GEUL* 274:171–200.

Jun, Youngchul. 2009. On 'i/ka' as a topic marker. *Discourse and Cognition* 16:217–238.

Jun, Youngchul. 2015. Topic, Focus, Contrast. In *The handbook of korean linguistics*, ed. Lucien Brown and Jaehoon Yeon, 179–195. Wiley-Blackwell.

Kang, Beom-mo. 2000. Nonhanguy uymi ywuhyengkwa tauyseng kiswul (Semantic type of arguments and the description of polysemy; written in Korean). In *Uymikwucouy phyosangkwa silhyen (the representationand realization of semantic structure)*.

Kang, Bosook. 2006. Some Peculiarities of Korean Kes Cleft Constructions. *Studia Linguistica* 60:251–281.

Karttunen, Lauri. 1971. Implicative Verbs. *Language* 47:340–358.

Kastner, Itamar. 2015. Factivity mirrors interpretation: The selectional requirements of presuppositional verbs. *Lingua* 164:156–188.

Kim, Ilkyu. 2015. Is Korean -(n)un a topic marker? On the nature of -(n)un and its relation to information structure. *Lingua* 154:87–109.

Kiparsky, Paul, and Carol Kiparsky. 1970. Fact. In *Progress in linguistics: A collection of papers*, ed. Manfred Bierwisch and Karl Erich Heidolph, 143–173. The Hague: Mouton.

Koopman, Hilda. 1983. ECP effects in main clauses. *Linguistic Inquiry* 14:346–350.

Kwon, Iksoo. 2012a. Please confirm what I inferred: On the Korean inferential-evidential marker *-napo-*. *Journal of Pragmatics* 44:958–969.

Kwon, Iksoo. 2012b. Viewpoints in the Korean verbal complex: Evidence, perception, assessment, and time. Doctoral Dissertation, UC Berkeley.

Ladusaw, William. 1994. Thetic and categorical, stage and individual, weak and strong. In *Semantics and Linguistic Theory*, volume 4, 220–229.

Lambrecht, Knud. 1994. *Information Structure and Sentence Form*. Cambridge University Press.

Lasnik, Howard, and Mamoru Saito. 1992. *Move alpha: Conditions on its application and output*. 22. The MIT press.

Leben, William. 1973. Suprasegmental phonology. Doctoral Dissertation, MIT.

Lee, Chungmin. 1976. Cases for psychological verbs in Korean. *Ene* 1:256–296.

Lee, Chungmin. 2003a. Contrastive Topic and proposition structure. In *Asymmetry in grammar*, 345–371.

Lee, Chungmin. 2003b. Contrastive Topic and/or Contrastive Focus. *Japanese/Korean Linguistics* 12:1–13.

Lee, Chungmin. 2017. Languages with non-factive knowledge verbs of 'know'. In *LENLS 14*.

Lee, Chungmin, Seungho Nam, and Beom-mo Kang. 1998. Lexical semantic structure for predicates in Korean. *Proceedings of ESSLLI-98* 1–15.

Lee, Hyeran. 2008. Left periphery in Korean: topicalization, focalization, and scrambling. *Studies in Generative Grammar* 18:137–171.

Levinson, Stephen. 1983. Cambridge University Press.

Li, Charles, and Sandra Thompson. 1981. *Mandarin Chinese: A functional reference grammar*. University of California Pres.

Maki, Hideki, Lizanne Kaiser, and Masao Ochi. 1999. Embedded topicalization in English and Japanese. *Lingua* 109:1–14.

Maling, Joan, and Annie Zaenen. 1977. TOP-position. *paper presented to the Annual Meeting of the Linguistic Society of America* .

McCarthy, John. 1986. OCP effects: Gemination and antigemination. *Linguistic Inquiry* 17:207–263.

Melvold, Janis. 1991. Factivity and definiteness. *MIT Working papers in linguistics* 15:97–117.

Mohanan, Tara. 1994. Case OCP: A constraint on word order in Hindi. In *Theoretical perspectives on word order in south asian languages*, ed. Miriam Butt, Tracy King, and Gillian Ramchand, 185–216. CSLI.

Moon, Gui-Sun. 1994. Topic Constructions in Korean. In *Explorations in generative grammar*, ed. Young-Sun Kim, Byung-Choon Lee, Kyoung-Jae Lee, Hyun-Kwon Yang, and Jong-Yurl Yoon, 435–454. Seoul: Hankuk Publishing Co.

Nichols, Lynn. 2001. *The syntax and morphosyntax of propositional attitude complementation*. Ms. Harvard University.

Obenauer, Hans-Georg. 1983. On the identification of empty categories. *The Linguistic Review* 4:153–202.

Obenauer, Hans-Georg. 1994. Aspects de la syntaxe A-barre: Effets d'intervention et mouvements des quantifieurs. Doctoral Dissertation, Université Paris 8.

Pak, Miok. 2008. Types of clauses and sentence end particles in Korean. *Korean Linguistics* 14:113–156.

Park, Yugyeong. 2015. A grammar of mood and clausal adjunction in Korean. Doctoral Dissertation, University of Delaware.

Portner, Paul. 2004. The Semantics of Imperatives within a Theory of Clause Types. *Semantics and Linguistic Theory* 16:235–252.

Prince, Ellen. 1981. Towards a taxonomy of given-new information. In *Radical pragmatics*, ed. Peter Cole, 223–256. Academic Press.

Reinhart, Tanya. 1981. Pragmatics and Linguistics: An Analysis of Sentence Topics. *Philosophica* 27:53–94.

Rizzi, Luigi. 1990. *Relativized Minimality*. The MIT Press.

Rizzi, Luigi. 1997. The Fine Structure of the Left Periphery. In *Elements of grammar*, ed. Liliane Haegeman, 281–337. Kluwer.

Rizzi, Luigi. 2001. Relativized Minimality Effects. In *The handbook of contemporary syntactic theory*, 89–110. Wiley-Blackwell.

Rizzi, Luigi. 2004. Locality and Left Periphery. In *Structures and beyond. the cartography of syntactic structures 3*, 223–251.

Rizzi, Luigi, and Giuliano Bocci. 2017. The left periphery of the clause: Primarily illustrated for Italian. In *The wiley blackwell companion to syntax*, ed. Martin Everaert and Henk van Riemsdijk, 1–30.

Roberts, Craige. 2011. Topics. In *Semantics: An international handbook of natural language meaning*, ed. Klaus von Heusinger, Claudia Maienborn, and Paul Portner, 1908–1934. De Gruyter.

Roussou, Anna. 2000. On the left periphery: modal particles and complementisers. *Journal of Greek linguistics* 1:65–94.

Saito, Mamoru. 2012. Sentence Types and the Japanese Right Periphery. In *Discourse and grammar: From sentence types to lexical categories*, ed. Günter Grewendorf and Thomas Zimmerman, 147–175. Berlin: Walter de Gruyter.

Shi, Dingxu. 2000. Topic and topic-comment constructions in Mandarin Chinese. *Language* 76:383–408.

Shyu, Shu-ing. 1995. The Syntax of Focus and Topic in Mandarin Chinese. Doctoral Dissertation, University of Southern California.

Shyu, Shu-ing. 2014. Topic and Focus. In *The handbook of chinese linguistics*, ed. James Huang, Audrey Li, and Andrew Simpson, 100–125. Wiley-Blackwell.

Sprouse, Jon. 2011. A validation of Amazon Mechanical Turk for the collection of acceptability judgments in linguistic theory. *Behavior research methods* 43:155–167.

Sprouse, Jon, Ivano Caponigro, Ciro Greco, and Carlo Cecchetto. 2016. Experimental syntax and the variation of island effects in English and Italian. *Natural Language & Linguistic Theory* 34:307–344.

Sprouse, Jon, Shin Fukuda, Hajime Ono, and Robert Kluender. 2011. Reverse Island Effects and the Backward Search for a Licensor in Multiple Wh-Questions. *Syntax* 14:179–203.

Stalnaker, Robert. 1975. Presuppositions. In *Contemporary research in philosophical logic and linguistic semantics*, 31–41.

Starke, Michal. 2001. Move dissolves into Merge: a theory of locality. Doctoral dissertation, University of Geneva.

Stowell, Timothy Angus. 1981. Origins of phrase structure. Doctoral Dissertation, Massachusetts Institute of Technology.

Strawson, Peter Frederick. 1964. Identifying reference and truth values. *Theoria* 30:96–118.

Tsai, Wei-Tien Dylan, ed. 2015. *The Cartography of Chinese Syntax*. Oxford University Press.

Vallduví, Enric. 1990. The Informational Component. Doctoral dissertation, University of Pennsylvania.

Vermeulen, Reiko. 2009. On the syntactic typology of topic marking: A comparative study of Japanese and Korean. In *Ucl working papers in linguistics*, UCL Working Papers in Linguistics, 335–363. London.

Vermeulen, Reiko. 2012. Word Order Variation and Information Structure in Japanese and Korean. In *The syntax of topic, focus and contrast*, ed. Ad Neeleman and Reiko Vermeulen, 77–118. De Gruyter.

Watanabe, Akira. 1993. Agr-based Case theory and its interaction with the A-bar system. Doctoral dissertation, MIT.

Appendix A: List of Figures and tables

List of Figures

List of Tables

Appendix B: Experimental materials for the main experiment

Token Set 1

[Nominative, With SFP]

Q: Ney-ka onul Ciho-ey tayhayse mwues-ul kiekhaynay-ss-e?

Q: 네가 오늘 지호에 대해서 무엇을 기억해냈어?

(What did you remember about Ciho today?)

A: Nay-ka onul Ciho-ka misin-ul mit-nun-ta-nun kes-ul kiekhaynay-ss-e.

A: 내가 오늘 지호가 미신을 믿는다는 것을 기억해냈어.

(I remembered today that Ciho is superstitious.)

[Nominative, Without SFP]

Q: Ney-ka onul Ciho-ey tayhayse mwues-ul kiekhaynay-ss-e?

Q: 네가 오늘 지호에 대해서 무엇을 기억해냈어?

(What did you remember about Ciho today?)

A: Nay-ka onul Ciho-ka misin-ul mit-nun kes-ul kiekhaynay-ss-e.

A: 내가 오늘 지호가 미신을 믿는 것을 기억해냈어.

(I remembered today that Ciho is superstitious.)

[Topic, With SFP]

Q: Ney-ka onul Ciho-ey tayhayse mwues-ul kiekhaynay-ss-e?

Q: 네가 오늘 지호에 대해서 무엇을 기억해냈어?

(What did you remember about Ciho today?)

A: Nay-ka onul Ciho-nun misin-ul mit-nun-ta-nun kes-ul kiekhaynay-ss-e.

A: 내가 오늘 지호는 미신을 믿는다는 것을 기억해냈어.

(I remembered today that Ciho (topicalized) is superstitious.)

[Topic, Without SFP]

Q: Ney-ka onul Ciho-ey tayhayse mwues-ul kiekhaynay-ss-e?

Q: 네가 오늘 지호에 대해서 무엇을 기억해냈어?

(What did you remember about Ciho today?)

A: Nay-ka onul Ciho-nun misin-ul mit-nun kes-ul kiekhaynay-ss-e.

A: 내가 오늘 지호는 미신을 믿는 것을 기억해냈어.

(I remembered today that Ciho (topicalized) is superstitious.)

Token Set 2

[Nominative, With SFP]

Q: Ney-ka akka Swua-ey tayhayse mwues-ul nwunchichay-ss-e?

Q: 네가 아까 수아에 대해서 무엇을 눈치챘어?

(What did you notice about Swua just a while ago?)

A: Nay-ka akka Swua-ka kohyang-ul kuliweha-n-ta-nun kes-ul nwunchichay-ss-e.

A: 내가 아까 수아가 고향을 그리워한다는 것을 눈치챘어.

(I noticed just a while ago that Swua misses her hometown.)

[Nominative, Without SFP]

Q: Ney-ka akka Swua-ey tayhayse mwues-ul nwunchichay-ss-e?

Q: 네가 아까 수아에 대해서 무엇을 눈치챘어?

(What did you notice about Swua just a while ago?)

A: Nay-ka akka Swua-ka kohyang-ul kuliweha-nun kes-ul nwunchichay-ss-e.

A: 내가 아까 수아가 고향을 그리워하는 것을 눈치챘어.

(I noticed just a while ago that Swua misses her hometown.)

[Topic, With SFP]

Q: Ney-ka akka Swua-ey tayhayse mwues-ul nwunchichay-ss-e?

Q: 네가 아까 수아에 대해서 무엇을 눈치챘어?

(What did you notice about Swua just a while ago?)

A: Nay-ka akka Swua-nun kohyang-ul kuliweha-n-ta-nun kes-ul nwunchichay-ss-e.

A: 내가 아까 수아는 고향을 그리워한다는 것을 눈치챘어.

(I noticed just a while ago that Swua (topicalized) misses her hometown.)

[Topic, Without SFP]

Q: Ney-ka akka Swua-ey tayhayse mwues-ul nwunchichay-ss-e?

Q: 네가 아까 수아에 대해서 무엇을 눈치챘어?

(What did you notice about Swua just a while ago?)

A: Nay-ka akka Swua-nun kohyang-ul kuliweha-nun kes-ul nwunchichay-ss-e.

A: 내가 아까 수아는 고향을 그리워하는 것을 눈치챘어.

(I noticed just a while ago that Swua (topicalized) misses her hometown.)

Token Set 3

[Nominative, With SFP]

Q: Ney-ka ecey Minho-ey tayhayse mwues-ul kippehay-ss-e?

Q: 네가 어제 민호에 대해서 무엇을 기뻐했어?

(What (thing) about Minho made you glad yesterday?)

A: Nay-ka ecey Minho-ka wuntong-ul culki-n-ta-nun kes-ul kippehay-ss-e.

A: 내가 어제 민호가 운동을 즐긴다는 것을 기뻐했어.

(I was glad yesterday that Minho enjoys sports.)

[Nominative, Without SFP]

Q: Ney-ka ecey Minho-ey tayhayse mwues-ul kippehay-ss-e?

Q: 네가 어제 민호에 대해서 무엇을 기뻐했어?

(What (thing) about Minho made you glad yesterday?)

A: Nay-ka ecey Minho-ka wuntong-ul culki-nun kes-ul kippehay-ss-e.

A: 내가 어제 민호가 운동을 즐기는 것을 기뻐했어.

(I was glad yesterday that Minho enjoys sports.)

[Topic, With SFP]

Q: Ney-ka ecey Minho-ey tayhayse mwues-ul kippehay-ss-e?

Q: 네가 어제 민호에 대해서 무엇을 기뻐했어?

(What (thing) about Minho made you glad yesterday?)

A: Nay-ka ecey Minho-nun wuntong-ul culki-n-ta-nun kes-ul kippehay-ss-e.

A: 내가 어제 민호는 운동을 즐긴다는 것을 기뻐했어.

(I was glad yesterday that Minho (topicalized) enjoys sports.)

[Topic, Without SFP]

Q: Ney-ka ecey Minho-ey tayhayse mwues-ul kippehay-ss-e?

Q: 네가 어제 민호에 대해서 무엇을 기뻐했어?

(What (thing) about Minho made you glad yesterday?)

A: Nay-ka ecey Minho-nun wuntong-ul culki-nun kes-ul kippehay-ss-e.

A: 내가 어제 민호는 운동을 즐기는 것을 기뻐했어.

(I was glad yesterday that Minho (topicalized) enjoys sports.)

Token Set 4

[Nominative, With SFP]

Q: Ney-ka ecey Sihyeni-ey tayhayse mwues-ul sulphehay-ss-e?

Q: 네가 어제 시현이에 대해서 무엇을 슬퍼했어?

(What (thing) about Sihyen made you sad yesterday?)

A: Nay-ka ecey Sihyeni-ka mwunhak-ul silheha-n-ta-nun kes-ul sulphehay-ss-e.

A: 내가 어제 시현이가 문학을 싫어한다는 것을 슬퍼했어.

(I was sad yesterday that Sihyen dislikes literature.)

[Nominative, Without SFP]

Q: Ney-ka ecey Sihyeni-ey tayhayse mwues-ul sulphehay-ss-e?

Q: 네가 어제 시현이에 대해서 무엇을 슬퍼했어?

(What (thing) about Sihyen made you sad yesterday?)

A: Nay-ka ecey Sihyeni-ka mwunhak-ul silheha-nun kes-ul sulphehay-ss-e.

A: 내가 어제 시현이가 문학을 싫어하는 것을 슬퍼했어.

(I was sad yesterday that Sihyen dislikes literature.)

[Topic, With SFP]

Q: Ney-ka ecey Sihyeni-ey tayhayse mwues-ul sulphehay-ss-e?

Q: 네가 어제 시현이에 대해서 무엇을 슬퍼했어?

(What (thing) about Sihyen made you sad yesterday?)

A: Nay-ka ecey Sihyeni-nun mwunhak-ul silheha-n-ta-nun kes-ul sulphehay-ss-e.

A: 내가 어제 시현이는 문학을 싫어한다는 것을 슬퍼했어.

(I was sad yesterday that Sihyen (topicalized) dislikes literature.)

[Topic, Without SFP]

Q: Ney-ka ecey Sihyeni-ey tayhayse mwues-ul sulphehay-ss-e?

Q: 네가 어제 시현이에 대해서 무엇을 슬퍼했어?

(What (thing) about Sihyen made you sad yesterday?)

A: Nay-ka ecey Sihyeni-nun mwunhak-ul silheha-nun kes-ul sulphehay-ss-e.

A: 내가 어제 시현이는 문학을 싫어하는 것을 슬퍼했어.

(I was sad yesterday that Sihyen (topicalized) dislikes literature.)

Token Set 5

[Nominative, With SFP]

Q: Ney-ka onul Hani-ey tayhayse mwues-ul kkaytal-ass-e?

Q: 네가 오늘 하니에 대해서 무엇을 깨달았어?

(What did you realize about Hani yesterday?)

A: Nay-ka onul Hani-ka kacok-ul salangha-n-ta-nun kes-ul kkaytal-ass-e.

A: 내가 오늘 하니가 가족을 사랑한다는 것을 깨달았어.

(I realized today that Hani loves (her) family.)

[Nominative, Without SFP]

Q: Ney-ka onul Hani-ey tayhayse mwues-ul kkaytal-ass-e?

Q: 네가 오늘 하니에 대해서 무엇을 깨달았어?

(What did you realize about Hani yesterday?)

A: Nay-ka onul Hani-ka kacok-ul salangha-nun kes-ul kkaytal-ass-e.

A: 내가 오늘 하니가 가족을 사랑하는 것을 깨달았어.

(I realized today that Hani loves (her) family.)

[Topic, With SFP]

Q: Ney-ka onul Hani-ey tayhayse mwues-ul kkaytal-ass-e?

Q: 네가 오늘 하니에 대해서 무엇을 깨달았어?

(What did you realize about Hani yesterday?)

A: Nay-ka onul Hani-nun kacok-ul salangha-n-ta-nun kes-ul kkaytal-ass-e.

A: 내가 오늘 하니는 가족을 사랑한다는 것을 깨달았어.

(I realized today that Hani (topicalized) loves (her) family.)

[Topic, Without SFP]

Q: Ney-ka onul Hani-ey tayhayse mwues-ul kkaytal-ass-e?

Q: 네가 오늘 하니에 대해서 무엇을 깨달았어?

(What did you realize about Hani yesterday?)

A: Nay-ka onul Hani-nun kacok-ul salangha-nun kes-ul kkaytal-ass-e.

A: 내가 오늘 하니는 가족을 사랑하는 것을 깨달았어.

(I realized today that Hani (topicalized) loves (her) family.)

Token Set 6

[Nominative, With SFP]

Q: Ney-ka akka Alami-ey tayhayse mwues-ul alanay-ss-e?

Q: 네가 아까 아람이에 대해서 무엇을 알아냈어?

(What did you find out about Alam just a while ago?)

A: Nay-ka akka Alami-ka kay-lul mwuseweha-n-ta-nun kes-ul alanay-ss-e.

A: 내가 아까 아람이가 개를 무서워한다는 것을 알아냈어.

(I found out just a while ago that Alam is afraid of dogs.)

[Nominative, Without SFP]

Q: Ney-ka akka Alami-ey tayhayse mwues-ul alanay-ss-e?

Q: 네가 아까 아람이에 대해서 무엇을 알아냈어?

(What did you find out about Alam just a while ago?)

A: Nay-ka akka Alami-ka kay-lul mwuseweha-nun kes-ul alanay-ss-e.

A: 내가 아까 아람이가 개를 무서워하는 것을 알아냈어.

(I found out just a while ago that Alam is afraid of dogs.)

[Topic, With SFP]

Q: Ney-ka akka Alami-ey tayhayse mwues-ul alanay-ss-e?

Q: 네가 아까 아람이에 대해서 무엇을 알아냈어?

(What did you find out about Alam just a while ago?)

A: Nay-ka akka Alami-nun kay-lul mwuseweha-n-ta-nun kes-ul alanay-ss-e.

A: 내가 아까 아람이는 개를 무서워한다는 것을 알아냈어.

(I found out just a while ago that Alam (topicalized) is afraid of dogs.)

[Topic, Without SFP]

Q: Ney-ka akka Alami-ey tayhayse mwues-ul alanay-ss-e?

Q: 네가 아까 아람이에 대해서 무엇을 알아냈어?

(What did you find out about Alam (topicalized) just a while ago?)

A: Nay-ka akka Alami-nun kay-lul mwuseweha-nun kes-ul alanay-ss-e.

A: 내가 아까 아람이는 개를 무서워하는 것을 알아냈어.

(I found out just a while ago that Alam (topicalized) is afraid of dogs.)

Token Set 7

[Nominative, With SFP]

Q: Ney-ka akka Tongswu-ey tayhayse mwues-ul nollawehay-ss-e?

Q: 네가 아까 동수에 대해서 무엇을 놀라워했어?

(What (thing) about Tongswu surprised you just a while ago?)

A: Nay-ka akka Tongswu-ka cwungkwuke-lul calha-n-ta-nun kes-ul nollawehay-ss-e.

A: 내가 아까 동수가 중국어를 잘한다는 것을 놀라워했어.

(I was surprised just a while ago that Tongswu is fluent in Chinese.)

[Nominative, Without SFP]

Q: Ney-ka akka Tongswu-ey tayhayse mwues-ul nollawehay-ss-e?

Q: 네가 아까 동수에 대해서 무엇을 놀라워했어?

(What (thing) about Tongswu surprised you just a while ago?)

A: Nay-ka akka Tongswu-ka cwungkwuke-lul calha-nun kes-ul nollawehay-ss-e.

A: 내가 아까 동수가 중국어를 잘하는 것을 놀라워했어.

(I was surprised just a while ago that Tongswu is fluent in Chinese.)

[Topic, With SFP]

Q: Ney-ka akka Tongswu-ey tayhayse mwues-ul nollawehay-ss-e?

Q: 네가 아까 동수에 대해서 무엇을 놀라워했어?

(What (thing) about Tongswu surprised you just a while ago?)

A: Nay-ka akka Tongswu-nun cwungkwuke-lul calha-n-ta-nun kes-ul nollawehay-ss-e.

A: 내가 아까 동수는 중국어를 잘한다는 것을 놀라워했어.

(I was surprised just a while ago that Tongswu (topicalized) is fluent in Chinese.)

[Topic, Without SFP]

Q: Ney-ka akka Tongswu-ey tayhayse mwues-ul nollawehay-ss-e?

Q: 네가 아까 동수에 대해서 무엇을 놀라워했어?

(What (thing) about Tongswu surprised you just a while ago?)

A: Nay-ka akka Tongswu-nun cwungkwuke-lul calha-nun kes-ul nollawehay-ss-e.

A: 내가 아까 동수는 중국어를 잘하는 것을 놀라워했어.

(I was surprised just a while ago that Tongswu (topicalized) is fluent in Chinese.)

Token Set 8

[Nominative, With SFP]

Q: Ney-ka ecey Sohuy-ey tayhayse mwues-ul anthakkawehay-ss-e?

Q: 네가 어제 소희에 대해서 무엇을 안타까워했어?

(What (thing) about Sohuy made you feel regretful yesterday?)

A: Nay-ka ecey Sohuy-ka chinkwutul-ul cilthwuha-n-ta-nun kes-ul anthakkawehay-ss-e.

A: 내가 어제 소희가 친구들을 질투한다는 것을 안타까워했어.

(I was regretful yesterday that Sohuy is jealous of her friends.)

[Nominative, Without SFP]

Q: Ney-ka ecey Sohuy-ey tayhayse mwues-ul anthakkawehay-ss-e?

Q: 네가 어제 소희에 대해서 무엇을 안타까워했어?

(What (thing) about Sohuy made you feel regretful yesterday?)

A: Nay-ka ecey Sohuy-ka chinkwutul-ul cilthwuha-nun kes-ul anthakkawehay-ss-e.

A: 내가 어제 소희가 친구들을 질투하는 것을 안타까워했어.

(I was regretful yesterday that Sohuy is jealous of her friends.)

[Topic, With SFP]

Q: Ney-ka ecey Sohuy-ey tayhayse mwues-ul anthakkawehay-ss-e?

Q: 네가 어제 소희에 대해서 무엇을 안타까워했어?

A: Nay-ka ecey Sohuy-nun chinkwutul-ul cilthwuha-n-ta-nun kes-ul anthakkawehay-ss-e.

(What (thing) about Sohuy made you feel regretful yesterday?)

A: 내가 어제 소희는 친구들을 질투한다는 것을 안타까워했어.

(I was regretful yesterday that Sohuy (topicalized) is jealous of her friends.)

[Topic, Without SFP]

Q: Ney-ka ecey Sohuy-ey tayhayse mwues-ul anthakkawehay-ss-e?

Q: 네가 어제 소희에 대해서 무엇을 안타까워했어?

(What (thing) about Sohuy made you feel regretful yesterday?)

A: Nay-ka ecey Sohuy-nun chinkwutul-ul cilthwuha-nun kes-ul anthakkawehay-ss-e.

A: 내가 어제 소희는 친구들을 질투하는 것을 안타까워했어.

(I was regretful yesterday that Sohuy (topicalized) is jealous of her friends.)

Appendix C: Experimental materials for the follow-up experiment[12]

Token Set 1: Adjunct (*because*) island

[Nominative-marked, Without SFP]

Q: Ney-ka akka Minhuy kwanlyenhayse iyakihay-ss-e? Mwusun iyaki-lul hay-ss-e?

Q: 네가 아까 민희 관련해서 이야기했어? 무슨 이야기를 했어?

(Did you talk about Minhuy just a while ago? What did you talk about?)

A: Nay-ka akka Minhuy-ka kecismalcayngi-iki ttaymwuney kyay-lul yokhay-ss-e.

A: 내가 아까 민희가 거짓말쟁이이기 때문에 걔를 욕했어.

(I criticized Minhuy because she is a liar.)

[Topic-marked, Without SFP]

Q: Ney-ka akka Cinkwu kwanlyenhayse iyakihay-ss-e? Mwusun iyaki-lul hay-ss-e?

Q: 네가 아까 진구 관련해서 이야기했어? 무슨 이야기를 했어?

(Did you talk about Cinkwu just a while ago? What did you talk about?)

A: Nay-ka akka Cinkwu-nun kecismalcayngi-iki ttaymwuney kyay-lul yokhay-ss-e.

A: 내가 아까 진구는 거짓말쟁이이기 때문에 걔를 욕했어.

(I criticized Cinkwu (topicalized) because he is a liar.)

Token Set 2: Adjunct (*because*) island

[Nominative-marked, Without SFP]

Q: Ney-ka onul Sengha kwanlyenhayse iyakihay-ss-e? Mwusun iyaki-lul hay-ss-e?

Q: 네가 오늘 성하 관련해서 이야기했어? 무슨 이야기를 했어?

(Did you talk about Sengha today? What did you talk about?)

A: Nay-ka onul Sengha-ka ppenppenha-ki ttaymwuney kyay-lul hyungpwa-ss-e.

A: 내가 오늘 성하가 뻔뻔하기 때문에 걔를 흉봤어.

(I talked badly of Sengha because she is shameless.)

[Topic-marked, Without SFP]

[12]The follow-up experiment was designed as a repeated measure, exposing participants to all stimuli. Being exposed to all conditions from the same token set may make it much easier to speculate what the research question is. As a minimal attempt to overcome this confound, different names were used for each item within the same token set. The lexical effect of name differences is expected to be insignificant.

Q: Ney-ka onul Mina kwanlyenhayse iyakihay-ss-e? Mwusun iyaki-lul hay-ss-e?
Q: 네가 오늘 민아 관련해서 이야기했어? 무슨 이야기를 했어?
(Did you talk about Mina today? What did you talk about?)
A: Nay-ka onul Mina-nun ppenppenha-ki ttaymwuney kyay-lul hyungpwa-ss-e.
A: 내가 오늘 민아는 뻔뻔하기 때문에 걔를 흉봤어.
(I talked badly of Mina (topicalized) because she is shameless.)

Token Set 3: Adjunct (*because*) island

[Nominative-marked, Without SFP]
Q: Ney-ka ecey Socini kwanlyenhayse iyakihay-ss-e? Mwusun iyaki-lul hay-ss-e?
Q: 네가 어제 소진이 관련해서 이야기했어? 무슨 이야기를 했어?
(Did you talk about Socin yesterday? What did you talk about?)
A: Nay-ka ecey Socini-ka sengsilha-ki ttaymwuney kyay-lul chingchanhay-ss-e.
A: 내가 어제 소진이가 성실하기 때문에 걔를 칭찬했어.
(I praised Socin because she is diligent.)

[Topic-marked, Without SFP]
Q: Ney-ka ecey Taycwuni kwanlyenhayse iyakihay-ss-e? Mwusun iyaki-lul hay-ss-e?
Q: 네가 어제 대준이 관련해서 이야기했어? 무슨 이야기를 했어?
(Did you talk about Taycwun yesterday? What did you talk about?)
A: Nay-ka ecey Taycwuni-nun sengsilha-ki ttaymwuney kyay-lul chingchanhay-ss-e.
A: 내가 어제 대준이는 성실하기 때문에 걔를 칭찬했어.
(I praised Taycwun (topicalized) because he is diligent.)

Token Set 4: Adjunct (*because*) island

[Nominative-marked, Without SFP]
Q: Ney-ka ecey Huymini kwanlyenhayse iyakihay-ss-e? Mwusun iyaki-lul hay-ss-e?
Q: 네가 어제 희민이 관련해서 이야기했어? 무슨 이야기를 했어?
(Did you talk about Huymin yesterday? What did you talk about?)
A: Nay-ka ecey Huymini-ka paysinca-iki ttaymwuney kyay-lul hemtamhay-ss-e.
A: 내가 어제 희민이가 배신자이기 때문에 걔를 험담했어.
(I slandered Huymin because he is a traitor.)

[Topic-marked, Without SFP]

Q: Ney-ka ecey Sanghwuni kwanlyenhayse iyakihay-ss-e? Mwusun iyaki-lul hay-ss-e?
Q: 네가 어제 상훈이 관련해서 이야기했어? 무슨 이야기를 했어?
(Did you talk about Sanghwun yesterday? What did you talk about?)
A: Nay-ka ecey Sanghwuni-nun paysinca-iki ttaymwuney kyay-lul hemtamhay-ss-e.
A: 내가 어제 상훈이는 배신자이기 때문에 걔를 험담했어.
(I slandered Sanghwun (topicalized) because he is a traitor)

Token Set 5: Relative clause island

[Nominative-marked, Without SFP]
Q: Ney-ka akka Minyongi kwanlyenhayse iyakihay-ss-e? Mwusun iyaki-lul hay-ss-e?
Q: 네가 아까 민용이 관련해서 이야기했어? 무슨 이야기를 했어?
(Did you talk about Minyong just a while ago? What did you talk about?)
A: Nay-ka akka Minyongi-ka sa-nun tongney-lul mwusihay-ss-e.
A: 내가 아까 민용이가 사는 동네를 무시했어.
(I made slighting remarks about the neighborhood where Minyong lives.)

[Topic-marked, Without SFP]
Q: Ney-ka akka Hyeyswu kwanlyenhayse iyakihay-ss-e? Mwusun iyaki-lul hay-ss-e?
Q: 네가 아까 혜수 관련해서 이야기했어? 무슨 이야기를 했어?
(Did you talk about Hyeswu just a while ago? What did you talk about?)
A: Nay-ka akka Hyeyswu-nun sa-nun tongney-lul mwusihay-ss-e.
A: 내가 아까 혜수는 사는 동네를 무시했어.
(I made slighting remarks about the neighborhood where Hyeswu (topicalized) lives.)

Token Set 6: Relative clause island

[Nominative-marked, Without SFP]
Q: Ney-ka akka Caywu kwanlyenhayse iyakihay-ss-e? Mwusun iyaki-lul hay-ss-e?
Q: 네가 아까 재우 관련해서 이야기했어? 무슨 이야기를 했어?
(Did you talk about Caywu just a while ago? What did you talk about?)
A: Nay-ka akka Caywu-ka tani-nun hoysa-lul pinanhay-ss-e.
A: 내가 아까 재우가 다니는 회사를 비난했어.
(I criticized the company that Caywu works for.)

[Topic-marked, Without SFP]

Q: Ney-ka akka Ihyeni kwanlyenhayse iyakihay-ss-e? Mwusun iyaki-lul hay-ss-e?
Q: 네가 아까 이현이 관련해서 이야기했어? 무슨 이야기를 했어?
(Did you talk about Ihyen just a while ago? What did you talk about?)
A: Nay-ka akka Ihyeni-nun tani-nun hoysa-lul pinanhay-ss-e.
A: 내가 아까 이현이는 다니는 회사를 비난했어.
(I criticized the company that Ihyen (topicalized) works for.)

Token Set 7: Relative clause island

[Nominative-marked, Without SFP]
Q: Ney-ka ecey Chanho kwanlyenhayse iyakihay-ss-e? Mwusun iyaki-lul hay-ss-e?
Q: 네가 어제 찬호 관련해서 이야기했어? 무슨 이야기를 했어?
(Did you talk about Chanho yesterday? What did you talk about?)
A: Nay-ka ecey Chanho-ka thayena-n nala-lul moyokhay-ss-e.
A: 내가 어제 찬호가 태어난 나라를 모욕했어.
(I insulted the country where Chanho was born.)

[Topic-marked, Without SFP]
Q: Ney-ka ecey Nami kwanlyenhayse iyakihay-ss-e? Mwusun iyaki-lul hay-ss-e?
Q: 네가 어제 나미 관련해서 이야기했어? 무슨 이야기를 했어?
(Did you talk about Nami yesterday? What did you talk about?)
A: Nay-ka ecey Nami-nun thayena-n nala-lul moyokhay-ss-e.
A: 내가 어제 나미는 태어난 나라를 모욕했어.
(I insulted the country where Nami (topicalized) was born.)

Token Set 8: Relative clause island

[Nominative-marked, Without SFP]
Q: Ney-ka onul Yongseni kwanlyenhayse iyakihay-ss-e? Mwusun iyaki-lul hay-ss-e?
Q: 네가 오늘 용선이 관련해서 이야기했어? 무슨 이야기를 했어?
(Did you talk about Yongsen today? What did you talk about?)
A: Nay-ka onul Yongseni-ka colepha-n hakkyo-lul chingchanhay-ss-e.
A: 내가 오늘 용선이가 졸업한 학교를 칭찬했어.
(I praised the school that Yongsen graduated from.)

[Topic-marked, Without SFP]

Q: Ney-ka onul Sangho kwanlyenhayse iyakihay-ss-e? Mwusun iyaki-lul hay-ss-e?
Q: 네가 오늘 상호 관련해서 이야기했어? 무슨 이야기를 했어?
(Did you talk about Sangho today? What did you talk about?)
A: Nay-ka onul Sangho-nun colepha-n hakkyo-lul chingchanhay-ss-e.
A: 내가 오늘 상호는 졸업한 학교를 칭찬했어.
(I praised the school that Sangho (topicalized) graduated from.)

Token Set 9: Complex NP island

[Nominative-marked, With SFP]
Q: Ney-ka onul Nahuy kwanlyenhayse iyakihay-ss-e? Mwusun iyaki-lul hay-ss-e?
Q: 네가 오늘 나희 관련해서 이야기했어? 무슨 이야기를 했어?
(Did you talk about Nahuy today? What did you talk about?)
A: Nay-ka onul Nahuy-ka misin-ul mitnun-ta-nun somwun-ul cenhay-ss-e.
A: 내가 오늘 나희가 미신을 믿는다는 소문을 전했어.
(I delivered the rumor that Nahuy is superstitious.)

[Topic-marked, With SFP
Q: Ney-ka onul Kangcwuni kwanlyenhayse iyakihay-ss-e? Mwusun iyaki-lul hay-ss-e?
Q: 네가 오늘 강준이 관련해서 이야기했어? 무슨 이야기를 했어?
(Did you talk about Kangcwun today? What did you talk about?)
A: Nay-ka onul Kangcwuni-nun misin-ul mitnun-ta-nun somwun-ul cenhay-ss-e.
A: 내가 오늘 강준이는 미신을 믿는다는 소문을 전했어.
(I delivered the rumor that Kangcwun (topicalized) is superstitious.)

Token Set 10: Complex NP island

[Nominative-marked, With SFP]
Q: Ney-ka akka Sunghyeni kwanlyenhayse iyakihay-ss-e? Mwusun iyaki-lul hay-ss-e?
Q: 네가 아까 승현이 관련해서 이야기했어? 무슨 이야기를 했어?
(Did you talk about Sunghyen just a while ago? What did you talk about?)
A: Nay-ka akka Sunghyeni-ka kacok-ul kuliwehan-ta-nun sasil-ul malhay-ss-e.
A: 내가 아까 승현이가 가족을 그리워한다는 사실을 말했어.
(I disclosed the fact that Sunghyen misses his family.)

[Topic-marked, With SFP]

Q: Ney-ka akka Hyeymini kwanlyenhayse iyakihay-ss-e? Mwusun iyaki-lul hay-ss-e?
Q: 네가 아까 혜민이 관련해서 이야기했어? 무슨 이야기를 했어?
(Did you talk about Hyemin just a while ago? What did you talk about?)
A: Nay-ka akka Hyeymini-nun kacok-ul kuliwehan-ta-nun sasil-ul malhay-ss-e.
A: 내가 아까 혜민이는 가족을 그리워한다는 사실을 말했어.
(I disclosed the fact that Hyemin (topicalized) misses her family.)

Token Set 11: Complex NP island

[Nominative-marked, With SFP]
Q: Ney-ka ecey Cwunghyeni kwanlyenhayse iyakihay-ss-e? Mwusun iyaki-lul hay-ss-e?
Q: 네가 어제 중현이 관련해서 이야기했어? 무슨 이야기를 했어?
(Did you talk about Cwunghyen yesterday? What did you talk about?)
A: Nay-ka ecey Cwunghyeni-ka kacoktul-ul salanghan-ta-nun kyellon-ul naylyesse.
A: 내가 어제 중현이가 가족들을 사랑한다는 결론을 내렸어.
(I made the conclusion that Cwunghyen loves his family.)

[Topic-marked, With SFP]
Q: Ney-ka ecey Yongho kwanlyenhayse iyakihay-ss-e? Mwusun iyaki-lul hay-ss-e?
Q: 네가 어제 용호 관련해서 이야기했어? 무슨 이야기를 했어?
(Did you talk about Yongho yesterday? What did you talk about?)
A: Nay-ka ecey Yongho-nun kacoktul-ul salanghan-ta-nun kyellon-ul naylyesse.
A: 내가 어제 용호는 가족들을 사랑한다는 결론을 내렸어.
(I made the conclusion that Yongho (topicalized) loves his family.)

Token Set 12: Complex NP island

[Nominative-marked, With SFP]
Q: Ney-ka akka Hyengmini kwanlyenhayse iyakihay-ss-e? Mwusun iyaki-lul hay-ss-e?
Q: 네가 아까 형민이 관련해서 이야기했어? 무슨 이야기를 했어?
(Did you talk about Hyengmin just a while ago? What did you talk about?)
A: Nay-ka akka Hyengmini-ka chinkwutul-ul cilthwuhan-ta-nun cwucang-ul ceysihay-ss-e.
A: 내가 아까 형민이가 친구들을 질투한다는 주장을 제시했어.
(I presented the argument that Hyengmin is jealous of his friends.)

[Topic-marked, With SFP]

Q: Ney-ka akka Kyenghuy kwanlyenhayse iyakihay-ss-e? Mwusun iyaki-lul hay-ss-e?
Q: 네가 아까 경희 관련해서 이야기했어? 무슨 이야기를 했어?
(Did you talk about Kyenghuy just a while ago? What did you talk about?)
A: Nay-ka akka Kyenghuy-nun chinkwutul-ul cilthwuhan-ta-nun cwucang-ul ceysihay-ss-e.
A: 내가 아까 경희는 친구들을 질투한다는 주장을 제시했어.
(I presented the argument that Kyenghuy (topicalized) is jealous of her friends.)

Token Set 13: *wh*-island

[Nominative-marked, With SFP]
Q: Ney-ka ecey Haycini kwanlyenhayse cilmwunhay-ss-e? Mwusun cilmwun-ul hay-ss-e?
Q: 네가 어제 해진이 관련해서 질문했어? 무슨 질문을 했어?
(Did you talk about Haycin yesterday? What did you talk about?)
A: Nay-ka ecey Haycini-ka nwukwu-lul salangha-nya-ko mwulepwa-ss-e.
A: 내가 어제 해진이가 누구를 사랑하냐고 물어봤어.
(I asked who Hyejin loves.)

[Topic-marked, With SFP]
Q: Ney-ka ecey Somi kwanlyenhayse cilmwunhay-ss-e? Mwusun cilmwun-ul hay-ss-e?
Q: 네가 어제 소미 관련해서 질문했어? 무슨 질문을 했어?
(Did you talk about Somi yesterday? What did you talk about?)
A: Nay-ka ecey Somi-nun nwukwu-lul salangha-nya-ko mwulepwa-ss-e.
A: 내가 어제 소미는 누구를 사랑하냐고 물어봤어.
(I asked who Somi (topicalized) loves.)

Token Set 14: *wh*-island

[Nominative-marked, With SFP]
Q: Ney-ka onul Cinsoli kwanlyenhayse cilmwunhay-ss-e? Mwusun cilmwun-ul hay-ss-e?
Q: 네가 오늘 진솔이 관련해서 질문했어? 무슨 질문을 했어?
(Did you talk about Cinsol today? What did you talk about?)
A: Nay-ka onul Cinsoli-ka mwues-ul mwuseweha-nya-ko cilmwunhay-ss-e.
A: 내가 오늘 진솔이가 무엇을 무서워하냐고 질문했어.
(I asked what Cinsol is afraid of.)

[Topic-marked, With SFP]

Q: Ney-ka onul Yongho kwanlyenhayse cilmwunhay-ss-e? Mwusun cilmwun-ul hay-ss-e?
Q: 네가 오늘 용호 관련해서 질문했어? 무슨 질문을 했어?
(Did you talk about Yongho today? What did you talk about?)
A: Nay-ka onul Yongho-nun mwues-ul mwuseweha-nya-ko cilmwunhay-ss-e.
A: 내가 오늘 용호는 무엇을 무서워하냐고 질문했어.
(I asked what Yongho (topicalized) is afraid of.)

Token Set 15: *wh*-island

[Nominative-marked, With SFP]
Q: Ney-ka akka Inswuki kwanlyenhayse cilmwunhay-ss-e? Mwusun cilmwun-ul hay-ss-e?
Q: 네가 아까 인숙이 관련해서 질문했어? 무슨 질문을 했어?
(Did you talk about Inswuk just a while ago? What did you talk about?)
A: Nay-ka akka Inswuki-ka eti-ey sa-nya-ko mwulepwa-ss-e.
A: 내가 아까 인숙이가 어디에 사냐고 물어봤어.
(I asked where Inswuk lives.)

[Topic-marked, With SFP]
Q: Ney-ka akka Kyenghwani kwanlyenhayse cilmwunhay-ss-e? Mwusun cilmwun-ul hay-ss-e?
Q: 네가 아까 경환이 관련해서 질문했어? 무슨 질문을 했어?
(Did you talk about Kyenghwan just a while ago? What did you talk about?)
A: Nay-ka akka Kyenghwani-nun eti-ey sa-nya-ko mwulepwa-ss-e.
A: 내가 아까 경환이는 어디에 사냐고 물어봤어.
(I asked where Kyenghwan (topicalized) lives.)

Token Set 16: *wh*-island

[Nominative-marked, With SFP]
Q: Ney-ka ecey Swucengi kwanlyenhayse cilmwunhay-ss-e? Mwusun cilmwun-ul hay-ss-e?
Q: 네가 어제 수정이 관련해서 질문했어? 무슨 질문을 했어
(Did you talk about Swuceng yesterday? What did you talk about?)
A: Nay-ka ecey Swucengi-ka eti-eyse ilha-nya-ko mwule-ss-e.
A: 내가 어제 수정이가 어디에서 일하냐고 물었어.
(I asked where Swuceng works.)

[Topic-marked, With SFP]

Q: Ney-ka ecey Sana kwanlyenhayse cilmwunhay-ss-e? Mwusun cilmwun-ul hay-ss-e?

Q: 네가 어제 사나 관련해서 질문했어? 무슨 질문을 했어?

(Did you talk about Sana yesterday? What did you talk about?)

A: Nay-ka ecey Sana-nun eti-eyse ilha-nya-ko mwule-ss-e.

A: 내가 어제 사나는 어디에서 일하냐고 물었어.

(I asked where Sana (topicalized) works.)

초록

사실성 동사의 보문절(사실성절)은 비사실성 동사의 보문절(비사실성절)과는 통사, 의미적으로 상당히 다른 성격을 보이는 것으로 알려져 있다. (Kiparsky and Kiparsky 1971) 의미적으로는 사실성절이 비사실성절과는 달리 주장의 의미를 결여하는 것으로 논의되어 왔다. (Hooper and Thompson 1973) 통사론의 측면에서는, Emonds (1976)에서 제시된 주절 이동 현상들이 사실성절 내에서는 인허되지 않는 것으로 관찰된다. 이 연구에서는 Emonds의 여러 주절 이동 현상들 중에서 특히 화제화에 초점을 맞춘다. 화제화는 Emonds 본인의 연구로부터 시작하여 여러 연구들에서 활발하게 연구된 현상이며. 화제화 역시 주절 이동 현상으로서 사실성절 내에서 인허되지 않는다. (Kiparsky and Kiparsky 1971, Watanabe 1993, Haegeman 2004)

화제화가 사실성절 내에서 인허되지 않는 현상을 설명한 이론은 크게 두 가지로 나누어 볼 수 있다. 첫째는 운용자를 활용한 접근이고, 둘째는 절 크기를 활용한 접근이다. 운용자 접근은 좀더 의미적인 설명을 제공한다. 사실성절이란 곧 어떤 사건에 대한 한정기술구와 유사하다는 Melvold (1991)의 주장을 기반으로 삼고 있으며, Watanabe (1993)와 Haegeman and Ürögdi (2010) 등에서 제시된다. 이 이론에 따르면 사실성절 외곽에는 '한정성' 또는 '사건' 등의 의미를 지닌 운용자가 위치하여 화제 인허를 방해한다.

반면 절 크기 접근은 좀 더 구조적인 설명 방식을 취한다. Haegeman (2004, 2006)이 Rizzi (1997)의 문장 외곽 이론을 적극적으로 수용하여 이러한 이론

을 주창하였다. Haegeman에 따르면 사실성절은 비사실성절에 비해 외곽의 기능범주들을 덜 투사한다. 즉 비사실성절이 투사하는 기능범주들을 사실성절은 투사하지 않는 것이다. 이 이론에 따르면 화제화가 인허되기 위해서는 비사실성절에만 존재하는 기능범주가 반드시 필요하기 때문에, 사실성절의 경우 화제를 인허하지 못한다.

한국어의 경우, 위 두 이론들에 새로운 경험적 문제를 제기한다. 여타 언어와 달리 한국어는 두가지의 사실성절 구조를 가지고 있기 때문이다. 시제소와 어말어미를 투사하는 '큰' 사실성절과, 이들을 투사하지 않는 '작은' 사실성절 두 가지가 있다. 이 연구에서는 한국어의 두 가지 사실성절 내부에서의 화제화에 대하여 기존의 이론들이 어떤 분석을 내릴 것인지, 그리고 이들의 분석이 과연 실제 한국어 화자들의 직관과 일치하는지 확인하고자 하였다.

이를 위하여 직관 판단 실험을 설계하였다. 사실성절과 화제화 등 통사-화용 접면의 언어 현상들의 경우 그 문법성의 판단이 극단적으로 갈리지 않기 때문에, 여러 화자들의 직관을 확인하여 각 이론의 예측과 화자들의 직관 판단이 유사한지 통계적으로 검증하고자 하였다. 실험 결과 운용자 이론과 절 크기 이론 모두 실험 참가자들의 직관 판단과 일치하는 예측을 내놓는 것으로 확인되었다. 더불어 추가 실험에서 한국어 섬 구조 내부에서의 화제화에 대한 직관을 관찰했을 때에는 절 크기 이론의 예측이 직관 측정 결과와 부분적으로 일치함을 확인할 수 있었다.

이 연구는 이전에 다루어지지 않은 언어 데이터를 가지고 기존의 이론들을

검증하고자 하였다는 데 의의가 있다. 더불어 절 크기 이론이 한국어의 사실성절과 섬 구조 등 통사적 구조를 설명하는 데에 도움이 된다는 점을 확인하였으며, 문말에 어말어미와 선어말어미 등 다양한 의존형태소가 분포하는 한국어의 특성상 절 크기 이론이 통사-화용 접면의 다른 현상들을 설명하는 데에도 기여하는 바가 있을 수 있음을 주장한다.

주요어: 사실성, 내포문, 화제화, 실험통사론, 통사-화용 접면부
학번: 2016-20061