

2018 Seoul Linguistics Forum

2018 서울언어학포럼

발표논문집

서울대학교

주최: 사단법인 한국언어학회, 한국언어정보학회, 서울대언어학과

후원: 서울대학교 <외국인저명학자 초청경비 지원사업>, <세계선도 중점학과 육성사업>

2018 Seoul Linguistics Forum

2018 서울언어학포럼

발표논문집

주제: Modality in Discourse

일시 2018년 12월 08일 (토) 09:30 ~ 18:50

장소 서울대학교 14동 대강당 (B101), 14동 204호

주최 사단법인 한국언어학회, 한국언어정보학회, 서울대언어학과

2018 Seoul Linguistics Forum (서울 언어학 포럼)은 서울대학교의 <외국인저명학자 초청경비 지원사업>과 <세계선도 중점학과 육성사업>의 도움을 받았습니다.

차 례

■ 주제 발표 <Modality in Discourse>

- The speaker–addressee relation in imperatives·····10

Paul Portner (Georgetown U.)

- Norms, discourse, and causation·····27

Aynat Rubinstein (Hebrew U. Jerusalem)

- Social Knowledge and Semantic Content ·····28

Elin McCready (Aoyama U.)

■ 개인 발표 1

- Epistemic questions in Korean and Japanese31
Takubo Yukinori (NINJAL)
- Psychological Origins and Levels of Deixis36
Nam Seungho (SNU)
- Korean evidential ‘-te’ and causal relations between situations47
Lim Dongsik (Hongik Univ.)
- Subjunctive complementizer in Korean: the interaction of inquisitiveness and nonveridicality48
강이름(고려대) · 윤수원(UTexas, Arlington)

■ 개인 발표 2

- ‘발화-인용구문’에 의한 목적 관계의 표현 — 유형론적 개관60
홍재성 (서울대학교)
- Creation of Richly Annotated Dialogue Corpus in Korean for Dialogue System: Making Out-of-Domain Dialogue Dataset from Large Scripts Data68
최지명 (연세대) · 김한샘 (연세대)
- 한국어의 ‘-음’과 ‘-기’에 대한 코퍼스 기반 연구74
조기현 (군산대) · 이용훈 (충남대)
- ‘기’ 구문의 정보 분할과 분열문80
김건희 (서원대)

■ 개인 발표 3

- Future Orientation and Korean Bouletic Verbs86
Choi Minkyung (SNU)
- **인식양태 구성들의 비교 연구**96
명정희 (서강대)
- Event Structure and Three Layers of Merge110
Davood Kasraie Kilakjani (SNU)
- To-infinitives vs. Gerunds in Spoken and Written English : A Case Study in the ICE-GB Corpus119
이규형 (한남대) · 이용훈 (충남대)

■ 개인 발표 4

- **영어 I-II 교과서와 대학수학능력시험 영어 영역에 나타난 등급화된 단어의 분포**128
이현우 (인하대)
- **한국어 분류사구와 측량 구문의 통사/의미론적 이질성: Ramchand(2008)의 동사구 분해 이론을 적용한 해명**138
김용민 (서울대)
- **술어와 논항의 구체성 정도 및 의미 타입 매치에 따른 맥락적 의미 변이 인식**149
임서현 (서울대)
- **한국어 동사의 구체성 정도 연구**159
박신영 (서울대)

2018 서울 언어학 포럼 프로그램

2018년 12월 8일 토요일 (서울대학교)		
09:10-09:30	등록 서울대학교 14동 대강당(B101) 로비	
09:30-09:40	개최식 서울대학교 14동 대강당(B101) [사회: 신용권(인천대)] 개최사 (사)한국언어학회 회장 최성호(충북대)	
기획주제 강좌 서울대 14동 대강당(B101) <Modality in Discourse> [좌장: 남승호(서울대)]		
09:40-10:45	Paul Portner (Georgetown U.) The speaker–addressee relation in imperatives	
10:45-11:50	Aynat Rubinstein (Hebrew U Jerusalem) Norms, discourse, and causation	
11:50-13:00	점심식사	
13:00-14:05	Elin McCready (Aoyama U.) Social Knowledge and Semantic Content	
14:05-14:30	휴식 / (사)한국언어학회 평의원회 · 임시 총회	
개인 발표 1 14동 대강당(B101) [좌장: 홍문표(성균관대)] 개인 발표 2 14동 204호 [좌장: 최동주(영남대)]		
14:30-15:00	Takubo Yukinori (NINJAL) Epistemic questions in Korean and Japanese	홍재성 (서울대) ‘발화-인용구문’에 의한 목적 관계의 표현 — 유형론적 개관
15:00-15:30	Nam Seungho (SNU) Psychological Origos and Levels of Deixis	최지명(연세대), 김한샘(연세대) Creation of Richly Annotated Dialogue Corpus in Korean for Dialogue System: Making Out-of-Domain Dialogue Dataset from Large Scripts Data
15:30-16:00	Lim Dongsik (Hongik Univ.) Korean evidential '-te' and causal relations between situations	조기현(군산대), 이용훈(충남대) 한국어의 ‘-고’와 ‘-기’에 대한 코퍼스 기반 연구
16:00-16:30	강아름(고려대), 윤수원(UTexas, Arlington) Subjunctive complementizer in Korean: the interaction of inquisitiveness and nonveridicality	김건희 (서원대) ‘기는’ 구문의 정보 분할과 분열문
16:30-16:40	휴식	

		개인 발표 3 14동 대강당(B101) [좌장: 김정수(경희대)]	개인 발표 4 14동 204호 [좌장: 이영주(서울여대)]
16:40-17:10		Choi Minkyung (SNU) Future Orientation and Korean Bouletic Verbs	이현우 (인하대) 영어 I-II 교과서와 대학수학능력시험 영어 영역에 나타난 등급화된 단어의 분포
17:10-17:40		명정희 (서강대) 인식양태 구성들의 비교 연구	김용민 (서울대) 한국어 분류사구와 측량 구문의 통사/의미론적 이질성: Ramchand(2008)의 동사구 분해 이론을 적용한 해명
17:40-18:10		Davood Kasraie Kilakjani (SNU) Event Structure and Three Layers of Merge	임서현 (서울대) 술어와 논항의 구체성 정도 및 의미 타입 매치에 따른 맥락적 의미 변이 인식
18:10-18:40		이규형(한남대)·이용훈(충남대) To-infinitives vs. Gerunds in Spoken and Written English: A Case Study in the ICE-GB Corpus	박신영 (서울대) 한국어 동사의 구체성 정도 연구
폐회식 14동 대강당(B101) [사회: 송상현(인천대)]			
18:40-18:50		폐회사 한국언어정보학회 회장 류병래(충남대)	

- 기획주제 강좌, 개인 발표 1·3 : 서울대학교 14동 대강당(B101)
- 개인 발표 2·4 : 서울대학교 14동 204호

[학술대회 안내]

- 일시: 2018년 12월 8일(토) 09:30 ~ 18:50
- 장소: 서울대학교 ① 서울대학교 14동 대강당(B101), ② 서울대학교 14동 204호
- 주제: Modality in Discourse
- 주최: 사단법인 한국언어학회, 한국언어정보학회, 서울대학교 언어학과

[등록비] ■ 일반: 20,000원 ■ 학생: 10,000원

- ※ 폐회식 후 저녁 만찬이 준비되어 있습니다.
- ※ 주차권이 필요하신 분은 말씀해 주시기 바랍니다.
- ※ 미리 참석 여부를 알려주시면 저희가 회장에서 선생님들을 맞을 준비를 하는데 큰 도움이 되오니 아래 이메일 주소로 참석 여부를 알려주시면 더욱 감사하겠습니다.
- ※ 연락주실 곳: linguisticsociety1@naver.com/010-9210-8480

■ 기획주제 강좌 ■

The speaker-addressee relation in imperatives*

Paul Portner
Georgetown University

December 8, 2018

1 The problem of embedded imperatives

Many languages do not allow imperatives to be embedded, while others allow embedding somewhat freely. Still other languages allow them in severely restricted circumstances.

1. No embedded imperatives: Italian

- (1) a. Chiamaci quando sei pronto.
call-IMP-us when are-2ND SG ready
'Call us when you're ready.'
- b. *Ha ordinato chiamaci quando sei pronto.
has ordered call-us when are-2ND SG ready
Intended meaning: 'He/she ordered that you call us when you're ready.'

2. Many embedded imperatives: Korean

- (2) Sensayngnim-i kyosil-ul chengsoha-la-ko
teacher-NOM classroom-ACC clean-IMP-COMP
cisiha/malssumha/pwuthakhasi-ess-ta.
order/say/request.HON-PAST-DEC
'The teacher ordered/said/requested (for us to) clean the classroom.'

3. Highly restricted embedded imperatives: English (with 'say')

- (3) John_i said call his_i mom. (Crnič and Trinh 2011, (6))

Previous work on embedded imperatives has focused more on the restrictions observed in one or several languages, and less on the variation that we see across languages (Kaufmann 2014 being a notable exception).

*The talk represents joint work with Miok Pak (George Washington University) and Raffaella Zanuttini (Yale University). A previous version was presented by the three collaborators at WAFL 14 at MIT.

Goal: Explain the contrast between the most extreme cases, i.e. languages that do not allow embedding at all versus those that seem to allow embedding with any semantically appropriate predicate.

- Our central intuition is that imperatives frequently encode a social/hierarchical relation between the speaker and the addressee. When they do, they cannot be embedded.
- Therefore, restrictions on embedding are not due to the illocutionary force or speech act type of imperatives, but rather to this ‘social’ meaning.

Clarifications:

- Many types of sentences that are used to perform a directive speech act can be embedded. For example, declaratives can be used to impose a requirement, and they can be embedded freely:

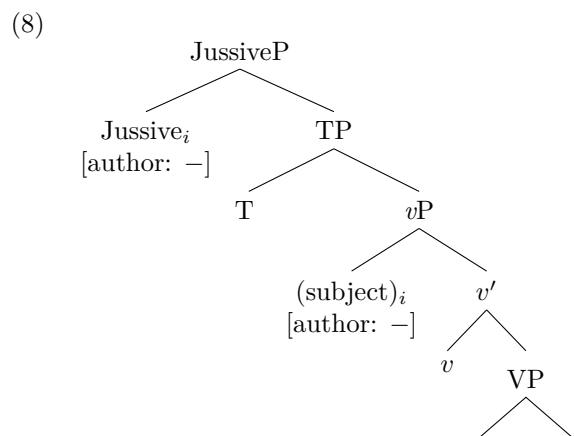
- (4) a. One must not park here.
 b. She said that one must not park here.

If we assume that the illocutionary force of (4a) is indirect, there’s no puzzle here for why embedding is possible in (4b).

- We focus on sentences that conventionally have a directive meaning. This can include a variety of syntactic forms. For example, in Italian we will include the canonical imperative with a verb in the imperative mood (5), so-called “polite imperatives” with a verb form in a different tense/mood combination like (6), and even infinitival directives like (7):

- (5) Chiama-ci quando sei pronto. (canonical imperative)
 call-IMP-us when are-2ND SG ready
 ‘Call us when you’re ready.’
- (6) Venga, si accomodi. (polite imperative)
 come-SUBJ.3S self make.comfortable-SUBJ.3S
 ‘Come in, make yourself comfortable.’
- (7) Rispettare l’ambiente! (infinitival directive)
 respect-INF the-environment
 ‘Respect the environment!’

- We assume, following Zanuttini et al. (2012) that imperatives (and other jussive clauses, i.e. exhortatives and promissives) contain a functional head, *JUSSIVE*. The Jussive head bears a person feature (2nd person, encoded as [-author]):



- JUSSIVE binds the subject and agrees in person with it, giving rise to the well-known person-restriction on imperative subjects. (This structure also leads to a property-type denotation that explains its force.)

Roadmap:

1. In section 2, we review our previous work which argued that the speaker-addressee relation is grammatically encoded as feature in a functional projection of the left periphery.
2. In section 3, we give the semantic/pragmatic analysis of these features, largely following our previous work but extending it in several ways.
3. In section 4, we provide our analysis of the restrictions on embedded imperatives in Korean and Italian.
4. In section 5, we summarize our proposals.

2 The syntax of the speaker-addressee relation

We aim to explain the unembeddability of imperatives (when they are in fact unembeddable) in terms of the fact that they encode information about the relation between the speaker and addressee in the conversation. So, we begin by presenting some ideas from a recent paper (Portner et al., to appear).

- The **speaker-addressee relation** is indicated in a wide variety of ways, including speech register, directness, and lexical choice (e.g. *please*).
 - Roughly, “politeness” (caveats)

- We aim for an analysis of **grammatical markers** of the speaker-addressee relation such as polite and familiar pronouns, vocatives, certain types of honorifics, allocutive agreement, politeness markers (e.g., Japanese *mas*) and speech style markers.
- As argued by Portner et al. (to appear), these fall into two categories:
 - CONTENT-ORIENTED MARKERS: Elements that mark the speaker’s relation to the *referent* of a noun phrase (who is, in certain cases, the addressee: polite/familiar second person pronouns).
 - UTTERANCE-ORIENTED MARKERS: Elements that mark the speaker’s relation to the *addressee*, who is not thereby introduced into the semantics as a referent.

Content-oriented markers of politeness. Italian marks the relation between the speaker and the addressee on second person pronouns. These pronouns are referred as ‘familiar’ and ‘polite’, or sometimes as *T* and *V* forms (after French *tu/vous*):

- (9) Gli infermieri **ti/Le** porteranno un bicchiere d’acqua.
 the nurses you.FAM/you.POL will-bring a glass of-water
 ‘The nurses will bring you a glass of water.’

These pronouns can readily be embedded:

- (10) Hanno detto che gli infermieri **ti/Le** porteranno un bicchiere d’acqua.
 have said that the nurses you.FAM/you.POL will-bring a glass of-water
 ‘They said that the nurses will bring you a glass of water.’

Other examples of content-oriented markers are the honorifics of Japanese and Korean (so-called referent or propositional honorifics).

Utterance-oriented markers of politeness. Korean marks the relation between speaker and addressee and the formality of the utterance context with a group of sentence final particles traditionally called ‘speech style particles’ (such as *formal*, *polite*, *plain*, and *intimate*). These particles also mark clause type. For example, the formal (deferential) speech style particle *-supnita* indicates that the addressee is older or socially higher than the speaker, the context of utterance is formal, and the sentence is a declarative.

- (11) Ecey pi-ka o-ass-**supnita**.
 yesterday rain-NOM come-PAST-DEC.FORMAL
 ‘It rained yesterday.’

These speech style particles cannot be embedded, with one apparent exception to be discussed later:

- (12) *Inho-ka [ecey pi-ka o-ass-**supnita**-ko] malhayss-supnita.
 Inho-NOM [yesterday rain-NOM come-PAST-DEC.FORMAL-COMP] said-DEC.FORMAL
 Intended meaning: ‘Inho said that it rained yesterday.’

Other examples of utterance-oriented markers are allocutive agreement, Japanese *-mas*, addressee agreement in Jingpo discussed by Zu (2018), and some of the discourse particles of West Flemish and Romanian as described by Haegeman and Hill (2013).

Little c. We have proposed that information about the speaker-addressee relation is encoded with the features [status] and [formal] in a functional category labeled “little c” (for *context*, Portner et al., to appear). The feature [formal] has values \pm and [status] has values that relate the speaker and addressee, e.g. $S \leq A$.

This projection is closely related to the Speech Act projections/layers of Speas and Tenny (2003), Miyagawa (2012), Haegeman and Hill (2013), Zu (2018), etc. We have argued that cP cannot be embedded and provided a semantic explanation for this restriction.

Evidence from Korean about the syntax of cP. We argue that Korean has two classes of particles within those traditionally called ‘speech style particles’:

- The **plain particles** do not mark a relation between speaker and addressee, and can only be used when the speaker doesn’t have a specific, direct interlocutor. They are commonly used in newspaper articles, personal diaries, professional journals, mottos and protest cries.

(13) Mwun taythonglyeng-i pwukhan-ul pangmuwnha-ess-**ta**.
 Moon president-NOM North.Korea-ACC visit-PAST-DEC.**plain**
 ‘The President Moon visited North Korea.’

- Speech style particles other than the plain particles mark the relation between speaker and addressee (cf. (11)), and can only be used when there is a specific, direct interlocutor. We refer to these particles simply as **speech style particles**, and we refer to an addressee who is a specific, direct interlocutor of the speaker as an INTERLOCUTOR-ADDRESSEE.
- In our previous work, we have argued that the particles traditionally associated with the “plain style” (*-ta*, *-nya/-ni*, *-(u)la/-ela*) actually represent two combinations of features. As (true) plain particles, they only indicate clause type, while as (plain) speech style particles, they also encode the speaker-addressee relation.

The two types of particles differ in embedding:

- Plain particles can be embedded, as shown in (14).

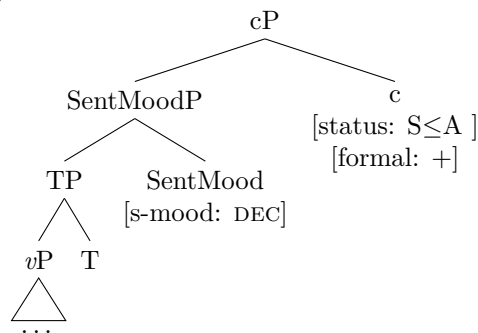
(14) [Mwun taythonglyeng-i pwukhan-ul pangmuwnha-ess-**ta**-ko] Yenhap
 Moon president-NOM North.Korea-ACC visit-PAST-DEC.**plain**-COMP Yenhap
 nyusu-ka onul potoha-ess-**ta**.
 News-NOM today report-PAST-DEC.PLAIN
 ‘The Yehap News reported today that President Moon visited North Korea.’

- Speech style particles cannot be embedded, as we saw in (12).

In Portner et al. (to appear), we argue that speech style particles (apart from the plain particles) realize both SentMood and *c*, and encode the relation between speaker and interlocutor-addressee with the [status] feature in *c*.

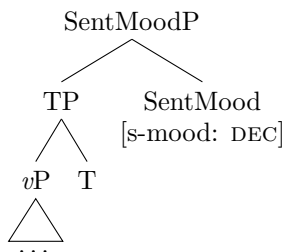
Hence, a sentence like (11) would contain the *c*P projection, as shown in (15):

(15) Speech style particles



In contrast, the plain speech style particles lack *c*P and realize only SentMood. Hence, the clausal structure of sentences such as (13) is as shown below:

(16) Plain particles



- *c*P cannot be embedded (we'll discuss why below). Hence sentences with speech style particles cannot be embedded (as in (12)).
- SentMoodP can be embedded. Hence sentences with plain particles can be embedded (as in (14)).

3 Semantics of the speaker-addressee relation

Our semantic analysis has four key ideas:

1. A component of the discourse context, the PARTICIPANT STRUCTURE, represents social/hierarchical relation between the speaker and interlocutor-addressee.
2. The [status] feature establishes the relation between participants in the participant structure.

3. The [formal] feature encodes information about which actual social relation is being represented.
4. Phrases which update the participant structure have a meaning in the “politeness dimension” π , and they cannot be embedded. Our analysis of these properties is closely related to prior work by Chierchia (1984), Potts and Kawahara (2004), Potts (2005), and McCready (2014).

The social meaning conveyed by [status] is “performative”, in that using it counts as an attempt to create the social relation it indicates. We can see this in the following examples:

- *Pretense*

- (17) Mom: Inho-ya, onul sihem cal poass-**ni**?
 Inho-VOC.INTIMATE, today test well done.INT.INTIMATE?
 ‘Inho, did you do well on the test today?’
- Inho: Ney, emma. 100 cem pat-ass-**eyo**.
 Yes, mom. 100 point receive-PAST-DEC.POLITE
 ‘Yes, mom. I got 100.’
- Mom: Wa! Cengmal? Cham calhayss-**eyo**!
 Wow! Really? Indeed well.done-DEC.POLITE
 ‘Wow! You did really well!’

- *Showing deference*

- (18) Boss: Kim tayli, onul hoyuy ilceng-i ettehkey toy-**pnikka**?
 Kim tayli.TITLE, today meeting schedule-NOM how become-INT.FORMAL
 ‘Assistant Kim, what is today’s meeting schedule?’
- Kim: Ney, thimcangnim. onul 3 si-ey makhething hoyuy-ka iss-**supnita**.
 Yes, boss. today 3 o’clock-at marketing meeting-NOM exist-DEC.FORMAL
 ‘Yes, boss. Today there is a marketing meeting at 3 o’clock.’

- *Shift to intimacy* (continuation of (18))

- (19) Boss: Kuntey onul cenyek-ey yaksok iss-**e**?
 By.the.way today evening-at plan have-INT.INTIMATE
 ‘By the way, do you have any plans for this evening?’
- Kim: Aniyoy, eps-**supnita**.
 No. not.have-DEC.FORMAL.
 ‘No. I don’t.’
- Boss: Yaksok epsumyen cenyek-ina kathi ha-**lkka**?
 plan have.not.if dinner-or.something together do-INT(SUGGEST).INTIMATE
 ‘If you have no plans, shall we have dinner or something?’

A formal model of the participant structure in discourse.

- (20) The **context** c is a pair $\langle P, cs \rangle$, where:
- i. the **participant structure** P is a triple $\langle J, O, h \rangle$, where
 - J is an n -tuple of individuals ($n \geq 1$), the **participants**,
 - O is an ordered set $\langle N, \prec \rangle$ with n members,
 - h is a function from J to subsets of N ;
 - ii. cs is the **context set**, a set of tuples $\langle x, y, t, w \rangle$

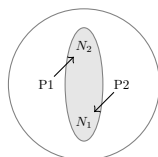


Figure 1: The participant structure

The status feature denotes a function of the same kind as h . For example (21a) is the semantics [status: S<A], and it goes into the politeness dimension π :

- (21) a. $\llbracket \text{[status: S<A]} \rrbracket^\pi = h : h(P_1) = \{N_1\}$ and $h(P_2) = \{N_2\}$
 b. $\llbracket \text{[status: S>A]} \rrbracket^\pi = h : h(P_1) = \{N_2\}$ and $h(P_2) = \{N_1\}$
 c. $\llbracket \text{[status: S≤A]} \rrbracket^\pi = h : h(P_1) = \{N_1\}$ and $h(P_2) = \{N_1, N_2\}$

When contextual update occurs, it sets the relation h in the participant structure to the one in the sentence's politeness dimension:

- (22) For utterance u of a declarative cP ϕ ,
- i. If the speaker of $u = P_1$, set h in the participant structure to $\llbracket \phi \rrbracket^\pi$:
 - $c + u = \langle \langle J_c, O_c, h_{c'} \rangle, cs' \rangle$, where
 - $h_{c'} = \llbracket \phi \rrbracket^\pi$,
 - $cs' = cs_c \cap \llbracket \phi \rrbracket^P$
 - ii. If the speaker of $u = P_2$, set h in the participant structure to $switch(\llbracket \phi \rrbracket^\pi)$:
 - $c + u = \langle \langle J_c, O_c, h_{c'} \rangle, cs' \rangle$, where
 - $h_{c'} = switch(\llbracket \phi \rrbracket^\pi)$,
 - $cs' = cs_c \cap switch(\llbracket \phi \rrbracket^P)$

See example in Figure 2.

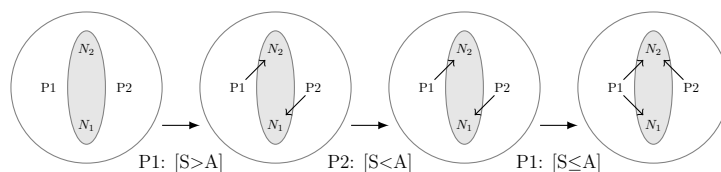
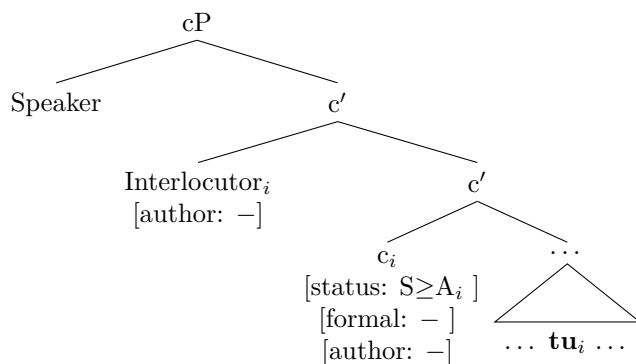


Figure 2: Updating the participant structure

Embeddability. In our recent work, we argue that a phrase with the type of “performative” meaning just described cannot be embedded. Thus, a cP with [status] cannot be embedded, but a clause lacking cP can be. To explain why content-oriented markers can be embedded, we propose that the [status] feature is interpreted only in the root c and, like person features (Baker 2008, Kratzer 2009), appears on the pronouns through agreement, as illustrated below for *tu* in Italian.

(23)



Formality and the interpretation of the participant structure. The hierarchical relation modeled by the participant structure has to be linked to some actual social relation between the participants. We achieve this with the following ALIGNMENT PRINCIPLE:

(24) **Alignment between participant structure and context set.** For every context c such that cs_c entails that there is a unique most salient social relation H involving the participants in the conversation, the ordering assigned to the participants in the participant structure in c is compatible with H .

In some cases, the alignment principle will simply reflect an salient relation already established in the context, while in others the use of form which marks the relation can cause a different hierarchy to become the most salient.

The alignment principle allows us to give a semantics for the [formal] feature.

1. Formality can be seen as a property of the salient social relation H . For example, a relation between boss and employee or between teacher and student would typically be [formal: +], while the relation between a parent and child would be [formal: -].

2. Formality and hierarchy can become closely related in context. The pronouns of Italian and other European languages can sometimes be better described as highlighting the formality or informality of the relation, where informality is typically associated with a less hierarchical situation (and hence hierarchy becomes moot). Similarly, the use of formal style in (18) is claimed to highlight the formality of the situation, while the hierarchy which is marked is largely *pro forma*.
3. Formality is closely related to other psychological and sociopragmatic concepts, such as affection, social distance, and autonomy. (By autonomy, I mean the extent to which the expectations associated with a relation can be negotiated by the participants themselves, as opposed to being fixed in the broader culture.) Ultimately, it might be useful to allow different components of the participant structure to “activate” in specific situations in the manner suggested by Eckert (2008).
4. It is difficult to know how to draw the line between what is encoded within the grammar, and what is outside the grammar but indirectly connected to it.

Social relation as a modal notion. It is useful to think of the social relation H in modal terms. When $H(a, b)$ represents the “boss-of” relation between a and b , we understand that a has certain powers over and responsibilities towards b , and vice versa. For example, if a issues a directive *Come into my office!* towards b , then b must come into a ’s office, and if b does not do so, a has the right to discipline b . But if H represents the “friend-of” relation, the powers and responsibilities are quite different.

1. These points suggest that $H(a, b)$ is a deontic conversational background (Kratzer 1981), a function from worlds to sets of propositions. Thus, if a is b ’s boss in w , then $H_{boss}(a, b)(w)$ entails ‘if a tells b to come into a ’s office, b comes into a ’s office.’
2. H is a formal social relation when it is only dependent on the roles of a and b as defined by society at large. For example, $H_{boss}(x, y)(w)$ gives the same set of requirements for x and y no matter who x and y are, provided that x is y ’s boss in w .
3. Different social relations have different properties as relations.
 - (a) “Boss-of” is antisymmetrical: If $H_{boss}(x, y)(w)$ is defined, then $H_{boss}(y, x)(w)$ is undefined.
 - (b) “Friend-of” is symmetrical: If $H_{friend}(x, y)(w)$ is defined, then $H_{friend}(y, x)(w)$ is also defined.
4. We hope to be able to define the relevant hierarchical [status] relations like $[S > A]$ and $[S = A]$ in terms of these types of properties of the social relations.

If the modal approach is correct, then the participant structure serves as an interface point between the categorical information marked by features in c and the properties of social relations (conversational backgrounds) that speakers need to evoke.

4 Analysis: embedded imperatives lack cP

Now we are going to give an analysis on the restrictions on embedded imperatives based on the ideas outlined above pertaining to speech style and jussives.

Our main ideas are captured in the following claims:

- As with the declarative clauses discussed above, imperatives can be embedded when they lack cP and cannot be embedded when they contain cP.
- In some languages but not others, the presence/absence of c in an imperative clause leads to a major difference in morphosyntax that affects whether the variant without c is called “imperative” in traditional descriptions.

What we will argue in this section is that both Korean and Italian have sentences where Jussive combines with c, and sentences where Jussive occurs without c, and that only the ones without c can be embedded. In some cases, we call the latter case an “imperative” (Korean plain style imperative), while in others it takes on a form that is traditionally not described this way (Italian infinitival directive).

There’s much more to do than can be done here to make these ideas completely precise in syntactic terms. Crucially, we need some further assumptions about pronouns:

- Overt second person pronouns (in Korean and Italian) represent both [author: –] and [status], and because of this they must refer to a specific addressee (i.e. an interlocutor of the speaker towards whom a [status] relation can be marked).
- A null pronoun can represent [author: –] in the absence of [status], and in such a case is used when the addressee is nonspecific/generic (i.e. not an interlocutor).

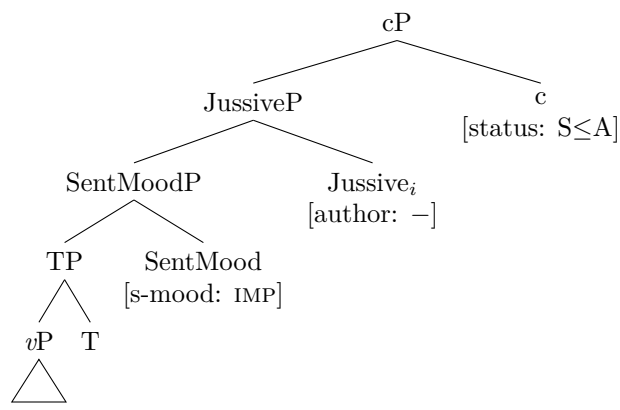
4.1 Korean

1. An imperative that marks speech style has both c and Jussive:

(25) Nayil wuli cip-ey o-si-eyo!
 tomorrow our house-to come-HON-IMP.POLITE
 ‘Come to my house tomorrow!’

In the structure for (25) shown in (26), c has a status feature [status: S \geq A] and Jussive has [author: –] and it binds and agrees with the subject.

(26) Imperative with a speech style particle in Korean



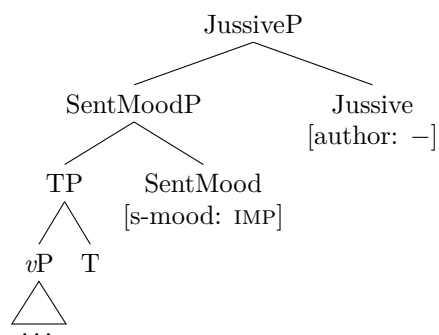
Because of presence of the [status] feature, this structure cannot be embedded:

- (27) *Inho-ka nayil cip-ey o-si-eyo-ko malhayss-supnita.
 Inho-NOM tomorrow house-to come-HON-IMP.POLITE-COMP said-DEC.FORMAL
 Intended meaning: ‘Inho said come to his house tomorrow.’

2. An imperative with a plain particle (normally used in writings, mottos or protest cries) contains Jussive but not c:

- (28) mithwu-ey ungtapha-**la**!
 #metoo-to respond-IMP.PLAIN
 ‘Respond to #metoo!’ (to the government or a company)

- (29) Imperative with a plain particle in Korean



Because it lacks c with the [status] feature, it is possible in both root and embedded clauses:

- (30) Inho-ka mithwu-ey ungtapha-**la**-ko malhayss-supnita.
 Inho-NOM #metoo-to respond-IMP.PLAIN-COMP said-DEC.FORMAL
 ‘Inho said respond to #metoo!’

3. As mentioned, we assume that a second person pronoun in Korean *ne* marks both [author: -] and [status] (cf. Portner et al., to appear).

Prediction:

- a second person pronoun is possible when *c* is present (with same [status] value).
- a second person pronoun is not possible in the absence of *c*.

This prediction is correct:

- (31) Ne(-nun) nayil wuli cip-ey o-**ala!**
 you(-TOP) tomorrow our house-to come-IMP.TRUEPLAIN
 ‘You come to my house tomorrow’

- (32) *Ne(-nun) mithwu-ey ungtapha-**la!**
 you(-TOP) #metoo-to respond-IMP.PLAINSTYLE
 Intended meaning: ‘You respond to #metoo!’ (to a specific person)

4.2 Italian

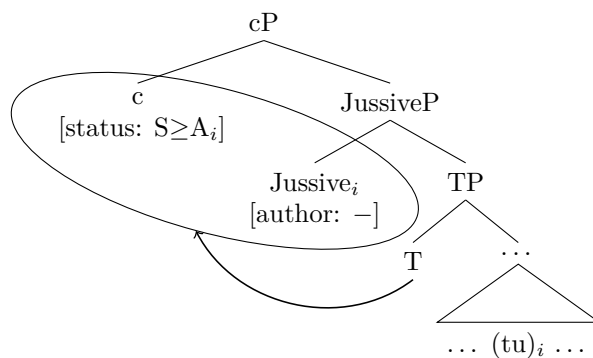
1. Two subtypes of imperatives mark the speaker-interlocutor relation:

The **canonical imperative** marks the ‘familiar’ (*tu*-form) relation:

- (33) Porta-mi l’acqua. (canonical imperative)
 bring.IMP-me the-water
 ‘Bring me the water.’

The structure contains both *c* and the Jussive head, which binds and agrees with the subject. We propose that in Italian canonical imperatives, J+c is a single head, hosting both [status: S_≥A] and [author: -]:

- (34)

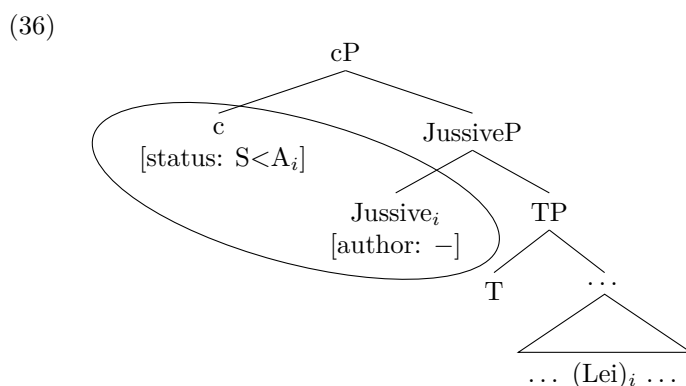


T raises to J+c, resulting in the canonical imperative verb form.

The **subjunctive polite imperative** marks the ‘polite’ (*Lei*-form) relation:

- (35) Mi porti l’acqua, per favore. (polite imperative)
 me bring.SUBJ the-water for favor
 ‘Would you bring me the water, please.’

This structure contains *c* with [status: S<A_{*i*}], and *J*, which brings in the [author: –] feature and binds the subject. As with the canonical imperative, *J* and *c* constitute a single head. However, in this case *T* does not raise and *c*+*J* selects the subjunctive form. (We don’t know why *T* does not raise; maybe something to do with the “fake” third person feature.)



The canonical imperative (33) cannot be embedded — as expected, given the presence of *c*:

- (37) *Ha ordinato porta-mi l’acqua.
 has ordered bring.IMP-me the-water
 Intended meaning: ‘He/she has ordered that you bring me the water.’

As for the polite imperative, it is hard to tell if it can be embedded. We assume it cannot, and that (38) contains a regular subjunctive without *Jussive* and *c*.

- (38) Ha ordinato che mi porti l’acqua.
 has ordered that me bring.SUBJ the-water
 ‘He/she ordered that you bring me the water.’

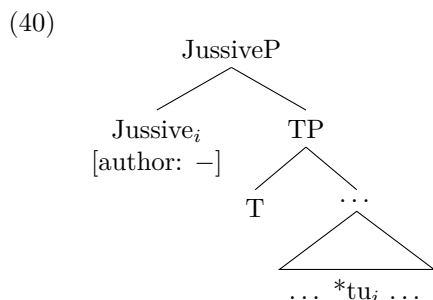
What is striking is that (38) loses the restriction that the subject of the subjunctive be the addressee: in this case, it could be another person as well.

(We can’t offer direct evidence on whether the polite imperative either can or cannot be embedded, since it would look just like the type of regular subjunctive we already know to be embeddable in the same position.)

2. An imperative-like meaning that is not directed towards a specific addressee is expressed with an infinitive. This form corresponds to the Korean nonspecific plain form imperative (28).

- (39) Rispettare l'ambiente. (infinitival directive)
 respect.INF the-environment
 'Respect the environment.'

This form contains Jussive but not *c*. We propose that *J* here selects infinitive *T*, which raises to *J*.



(In principle, we expect that the form in (40) is embeddable. It's difficult to tell, though, since they have the same PF form as a regular control infinitive. Examples would have a form like *She told him to leave*, where the embedded subject must be the addressee of the reported speech act, but such examples can be handled by control theory.)

3. As with the Korean (32), the form without *c* (the infinitival imperative) cannot contain a second person pronoun:

- (41) Dopo l'esame, restituire le calcolatrici agli insegnanti.
 after the-exam return the calculators to.the teachers
 'After the exam, return the calculators to the teachers.'
- (42) *Dopo l'esame, restituire le calcolatrici ai tuoi insegnanti.
 after the-exam return the calculators to.the your teachers
 Intended meaning: 'After the exam, return the calculators to your teachers.'
- (43) Dopo l'esame, restituire le calcolatrici ai propri insegnanti.
 after the-exam return the calculators to.the own teachers
 'After the exam, return the calculators to one's own teachers.'

This restriction follows on the assumption that both polite and familiar second person pronouns mark [status], which enters the derivation in *c*.

5 Summary

1. We pursue a new idea about why there are restrictions on embedding imperatives. We propose that imperatives often encode expressive information about the speaker-interlocutor relation, and that sentences with such meaning cannot be embedded. (It's not their directive illocutionary force which prevents embedding, contra Han 1998, etc.)
2. We discussed the case of Korean, and proposed that imperatives can be embedded because the language makes available an imperative form that does not mark the relation between the speaker and interlocutor-addressee.
3. We explain why imperatives cannot be embedded in languages like Italian by tying the canonical imperative form to a feature that does mark such meaning.
4. We drew a connection between the Korean plain form and the Italian infinitival directive. These sentences have several properties that follow from our analysis: their "generic addressee" meaning, a prohibition on second person pronouns, and embeddability.

References

- Baker, M. C. (2008). *The Syntax of Agreement and Concord*. Cambridge University Press, Cambridge.
- Chierchia, G. (1984). *Topics in the Syntax and Semantics of Infinitives and Gerunds*. PhD thesis, University of Massachusetts, Amherst.
- Crnić, L. and Trinh, T. (2011). Embedding imperatives. In Lima, S., Mullin, K., and Smith, B., editors, *Proceedings of NELS 39*, volume 1, Amherst. GLSA, Univ. of Massachusetts.
- Eckert, P. (2008). Variation and the indexical field. *Journal of Sociolinguistics*, 12(4):453–476.
- Haegeman, L. and Hill, V. (2013). The syntacticization of discourse. In Folli, R., Sevdali, C., and Truswell, R., editors, *Syntax and its limits*, pages 370–390. Oxford University Press, Oxford.
- Han, C.-H. (1998). *The syntax and semantics of Imperatives and Related Constructions*. PhD thesis, University of Pennsylvania.
- Kaufmann, M. (2014). Embedded imperatives across languages: Too rare to expect, too frequent to ban. Handout, Colloquium, Department of Linguistics, SUNY Stony Brook.
- Kratzer, A. (1981). The notional category of modality. In Eikmeyer, H.-J. and Rieser, H., editors, *Words, Worlds, and Contexts*, pages 38–74. de Gruyter, Berlin.
- Kratzer, A. (2009). Making a pronoun: Fake indexicals as windows into the properties of pronouns. *Linguistic Inquiry*, 40(2):187–237.
- McCready, E. (2014). A semantics for honorifics with reference to Thai. In Aroonmanakun, W., Boonkwan, P., and Supnithi, T., editors, *28th Pacific Asia Conference on Language, Information and Computing (PACLIC 28)*, pages 503–512, Stroudsburg, PA. Association for Computational Linguistics (ACL).
- Miyagawa, S. (2012). Agreements that occur mainly in the main clause. In Aelbrecht, L., Haegeman, L., and Nye, R., editors, *Main Clause Phenomena: New Horizons*, pages 79–112. John Benjamins, Amsterdam/Philadelphia.

- Portner, P., Pak, M., and Zanuttini, R. (to appear). The speaker-addressee relation at the syntax-semantics interface. *Language*.
- Potts, C. (2005). *The Logic of Conventional Implicature*. Oxford University Press, Oxford.
- Potts, C. and Kawahara, S. (2004). Japanese honorifics as emotive definite descriptions. In Watanabe, K. and Young, R. B., editors, *Proceedings of Semantics and Linguistic Theory 14*, pages 235–254. CLC Publications, Ithaca, NY.
- Speas, M. and Tenny, C. L. (2003). Configurational properties of point of view roles. In Sciullo, A. M. D., editor, *Asymmetry in grammar*, pages 315–344. John Benjamins, Amsterdam/Philadelphia.
- Zanuttini, R., Pak, M., and Portner, P. (2012). A syntactic analysis of interpretive restrictions on imperative, promissive and exhortative subjects. *Natural Language and Linguistic Theory*, 30(4):1231–1274.
- Zu, V. (2018). *Discourse Participants and the Structural Representation of the Context*. PhD thesis, New York University, New York, NY.

Norms, discourse, and causation

Aynat Rubinstein (Hebrew U. Jerusalem)

Causal judgment shows sensitivity to norms in the sense that what is morally good, legally required, or epistemically expected affects our intuitions about what count as possible and impossible causes. Although it shines through in our intuitions about causative statements, the normative aspect of causation has been argued to have dubious theoretical status. In this talk, I explore connections between degrees of modal necessity and causality that may offer a way to ground the normative component of causal language in a more solid foundation, relying on the status of norms in discourse.

Social Knowledge and Semantic Content

Elin McCready

In linguistic discussions of honorification, the main topic tends to be honorific morphology (as in Japanese) or, occasionally, particles (as in Thai). Phenomena like titles and the use of information about profession and social role is less discussed. This talk focuses on what I call role honorifics, where information about prestige and social position are used to perform (anti)honorification. I argue for an expressive treatment, where titles and pronominally used nominals introduce information relevant to social roles to not-at-issue content; this content is then used to derive honorific effects via alterations in the current register as a result of nonmonotonic inferences about language use. The analysis is then applied to the role of normative gender in pronominal interpretation.

■ 개인 발표 1 ■

Epistemic questions in Korean and Japanese

Yukinori Takubo
NINJAL
ytakubo@ninjal.ac.jp

1.1 A 'future' marker *-I kes-i-* under epistemic reading cannot be used in true questions (Korean)

- (1) a. Nayil na-nun Seoul-ey **ka-l kes-i**-pni-ta. (Volitional)
tomorrow I-TOP Seoul-to go-I kes-i-HON-IND
"I will go to Seoul tomorrow."
b. Nayil-un pi-ka **o-l kes-i**-pni-ta. (Epistemic)
tomorrow-TOP rain-NOM come-I kes-i-HON-IND
"(I think) it will rain tomorrow."
- (2) a. Nayil tangsın-un **o-l kes-i**-pni-kka?
(Volitional) tomorrow you-TOP come-I kes-i-HON-Q
"Are you coming tomorrow?"
b. * Nayil-un pi-ka **o-l kes-i**-pni-kka? (Epistemic)
tomorrow-TOP rain-NOM come-I kes-i-HON-Q
"(Do you think) it will rain tomorrow?"

(3) The speaker greets a farmer who is looking up at the sky;

- a. Nayil pi-kao-**keyss**-supni-kka?
tomorrow rain-NOM come-keyss-HON-Q
"Will it rain tomorrow?"
b. Nayil pi-kao-l **kka**-yo?
tomorrow rain-NOM come-I Q-HON
"Will it rain tomorrow?"

1.2. Japanese hazu (≈must) cannot be used in true questions either

- (4) a. Zimen-ga nureteiru **hazu**-desu.
ground-NOM be.wet hazu-COP.HON
"The ground must be wet."
b. Yamada-wa inal **hazu**-desu.
Yamada-TOP be.absent hazu-COP.HON
"Yamada must be absent."
- (5) a. ?? Zimen-ga nureteiru **hazu**-desu-ka?
ground-NOM be.wet hazu-COP.HOM-Q
"The ground must be wet."
b. ?? Yamada-wa inal **hazu**-desu-ka?
Yamada-TOP be.absent hazu-COP.HON-Q
"Yamada must be absent."

1.3. Point of view shift in questions

- (6) a. He will be admitted to SNU. ('I think he will be...')
b. Will he be admitted to SNU? ('Do you think he will be...')

Point of view shift in **will**

Affirmative sentence
compatible with the knowledge of the speaker

Interrogative sentence
compatible with the knowledge of the hearer

- (4) a. Zimen-ga nureteiru **hazu**-desu.
ground-NOM be.wet hazu-COP.HON
"The ground must be wet."
(5) a. ?? Zimen-ga nureteiru **hazu**-desu-ka?
ground-NOM be.wet hazu-COP.HON-Q
"The ground must be wet Q"

Takubo and Kim (2009) attempted to explain the contrast in (4)–(5) in terms of the point of view shift or the lack thereof, namely, by stipulating that epistemic necessity modals do not allow deictic shift. Below we will take a close look at **hazu-da** and give a principled account for their stipulation based on the lexical semantics of **hazu-da**.

2. Japanese hazu

2.1. “q-hazu-da” signals that q is a logical consequence of some sort Takubo (2001, 2009)

(7) “the ground is wet-*hazu-da*” (=4a)

1. if it rains, the ground is wet
2. it rained
3. the ground is wet

(8) “Yamada is absent-*hazu-da*” (=4b)

1. if x’s travel document has been submitted, x is absent
2. Yamada’s travel document has been submitted
3. Yamada is absent

“q-*hazu-da*” presupposes a certain set of premises *P* that make *q* true.
In the case of modus ponens $P = \{ \text{if } p \text{ then } q \}, \{ p \}$
cf. Conversational Background ((given) “what we know”) (Kratzer 1991)

2.1. “q-hazu-da” signals that q is a logical consequence of some sort

Also in cases where no causation is involved

(9) “Tanaka has got to Tokyo now-*hazu-da*”

1. Yamada left Tokyo to Kyoto at 8am and Tanaka left Kyoto to Tokyo at 8am
2. Yamada has just got to Kyoto now
3. Tanaka has got to Tokyo now

“q-*hazu*” presupposes some set of propositions *P* that make *q* true.

2.2. “q-hazu?” is ruled out because of self contradiction

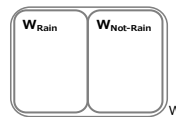
Ingredient 1. “the ground is wet-hazu”

Suppose we all know: if it rains, then the ground is wet
premise consequence

1. if it rains, the ground is wet ... we all know
2. it rained ... premise observed
3. the ground is wet ... consequence drawn
→ “the ground is wet-*hazu-da*”
(under normal circumstances) (Kratzer 1981, 1991, 2012)

Ingredient 2. Yes/No-questions: partition (Groenendijk & Stokoff 1984)

(10) Is it raining now?



W_{Rain} : a set of possible worlds in which it is raining now
 $W_{\text{Not-Rain}}$: a set of possible worlds in which it is not raining now
 W : the set of all possible worlds

(11) Are we in W_{Rain} or $W_{\text{Not-Rain}}$?

2.3. “q-hazu?” is ruled out because of self contradiction

“the ground is wet-*hazu?*” ... (a)

1. if it rains, the ground is wet ... (b) we all know
2. it rained ... (c) premise observed
3. the ground is wet ... (d) consequence drawn

→ “the ground is wet-*hazu?*”
(under normal circumstances) (Kratzer 1981, 1991, 2012)

1. “q-*hazu?*” (a) presupposes a set of propositions (b)(c) that make *q* true-*hazu* (P-hole)
2. “q-*hazu?*” (a) → “Are we in W_q or in $W_{\neg q}$,” ... yes/no-Q
3. P (b)(c) entails *q* (d). 2 presupposes you don’t know the answer.

Therefore:

- “q-*hazu?*” is not a legitimate question

3. Japanese noda

3.1. noda (nominal complementizer+copula) saves “q-hazu?”

(12) a. ?? Zimen-ga nureteiru **hazu**-desu-ka?
ground-NOM be.wet hazu-COP.HON-Q
“The ground must be wet.”

b. ?? Yamada-wa inai **hazu**-desu-ka?
Yamada-TOP be.absent hazu-COP.HON-Q
“Yamada must be absent.” (=2)

(13) a. Zimen-ga nureteiru **hazu-na-no**desu-ka?
ground-NOM be.wet hazu-COP.ADNOM-noda.HON-Q
“The ground must be wet.”

b. Yamada-wa inai **hazu-na-no**desu-ka?
Yamada-TOP be.absent hazu-COP.ADNOM-noda.HON-Q
“Yamada must be absent.”

3.2. noda can trigger abduction (Takubo 2009)

Suppose we all know: if it rains, then the ground is wet
 premise consequence

Deduction: By observing the premise(s), you draw the consequence.
 if it rains, the ground is wet ... we all know
 it rained ... premise observed
 the ground is wet ... consequence drawn
 → "the ground is wet-hazu"

Abduction: By observing the consequence, you hypothesize (one of) the premise(s).
 if it rains, the ground is wet ... we all know
 the ground is wet ... consequence observed
 it rained ... premise hypothesized
 → "It rained-noda"

3.3. "q-hazu-(na)-noda?" is a question about (one of) the premise(s)

By adding **noda** to "q-hazu?", you can ask what q presupposes.
 Someone with very few experiences of rain says:

(15) Zimen-ga nureteiru **hazu-na-nodesu-ka?**
 ground-NOM wet hazu-COP.ADNM-noda.HON-Q
 "Is it true that (in this country if it rains (usually))
 the ground is wet?" meaning "Is that what you would expect?"

if it rains, the ground is wet ← asking this part
it rained
 the ground is wet

3.3. "q-hazu-(na)-noda?" is a question about (one of) the premise(s)

By adding **noda** to "q-hazu?", you can ask what q presupposes.
 Someone who doesn't know what's going on in the admin office says:

(16) Yamada-wa inai **hazu-na-nodesu-ka?**
 Yamada-TOP be.absent hazu-COP.ADNM-noda.HON-Q
 "Is it true that (Yamada's travel document has been submitted
 and thus) he is absent?"

if x's travel document has been submitted, x is absent
 Yamada's travel document has been submitted ← asking this part
 Yamada is absent

3.3. "q-hazu-(na)-no?" is a question about (one of) the premise(s)

"Yamada is absent-hazu-na-no?"
 If x's travel document has been submitted, x is absent
 Yamada's travel document has been submitted ← asking this part
 Yamada is absent

"q-hazu-(na)-no?" is a question about one of the premises
 If p, then q, p|=q
 → "p or ¬p, which one is it?"

Therefore:
 "q-hazu-na-no?" is a legitimate question
 "wide scope interpretation of **noda**"

4. Korean -l kes-i-

4.1. Korean "q-l kes-i-" patterns with "q-hazu"

-l kes-i- when used as epistemic modal cannot be used as true questions.

(17) a. Nayil na-nun Seoul-ey **ka-l kes-i-pni-ta.** (Volitional)
 tomorrow I-TOP Seoul-to go-I kes-i-HON-IND
 "I will go to Seoul tomorrow."
 b. Nayil-un pi-ka **o-l kes-i-pni-ta.** (Epistemic)
 tomorrow-TOP rain-NOM come-I kes-i-HON-IND
 "I (think) it will rain tomorrow."

(18) a. Nayil tangsin-un **o-l kes-i-pni-kka?** (Volitional)
 tomorrow-you-TOP come-I kes-i-HON-Q
 "Are you coming tomorrow?"
 b. Nayil-un pi-ka **o-l kes-i-pni-kka?** (Epistemic)
 tomorrow-TOP rain-NOM come-I kes-i-HON-Q
 "(Do you think) it will rain tomorrow?"

4.1. Korean "q-l kes-i-" patterns with "q-hazu"

"q-l kes-i-" expresses 'logical consequences' just like "q-hazu" and presupposes that there are premises from which q can be deduced.

(19) Pi-ga **o-l kes-i-pni-ta.**
 rain-NOM come-I kes-i-HON-IND
 "It will rain."

if there are halos around the moon, it will rain
there are halos around the moon
 it will rain
 "q-l kes-i-" in this use cannot be questioned because it is interpreted that you are questioning what you just deduced, i.e. you are denying your presuppositions.

4.2. “q-l kes-i-“ cannot be made into a meta-question(i.e. a question about the premises)

(20) * Pi-ka o-l kes-i-pni-kka?
rain-NOM come-I kes-i-HON-Q
“(Do you think) it will rain?”

To ask a question about the premises in Japanese, **noda** is used to expand the scope to include premises.

This is impossible in Korean because of the morphological make-up of **-l kes-i-**.

4.3. Meta-questions in Korean can be formed by attaching -nun kes-i- for present and -n kes-i- for past eventualities.

(21) a. Pi-kao-nun kes-i-pni-kka?
rain-NOM come-PRES.ADNOM kes-i-HON-Q
“You have an umbrella with you (because it will rain)?”
b. Pi-kao-n kes-i-pni-kka?
rain-NOM come-PAST.ADNOM kes-i-HON-Q
“You are all wet (because it rained)?”

4.3. Meta-questions in Korean can be formed by attaching -nun kes-i- for present and -n kes-i- for past eventualities.

Morphologically **-l kes-i-** forms a paradigm with these meta-markers even though it is fully grammaticalized to become a modal auxiliary. So it cannot be followed by these meta-markers. (Notice that for copula **-l** the present adnominal form is **-n** and past adnominal form is **-ten**.)

(22) a. * Pi-ka o-l kes-i-n kes-i-pni-kka?
rain-NOM come-I kes-i-PRES.ADNOM kes-i-HON-Q
Int. “Is it because it will rain(that you have an umbrella with you)?”
b. * Pi-ka o-l kes-i-ten kes-i-pni-kka?
rain-NOM come-I kes-i-PAST.ADNOM kes-i-HON-Q
Int. “Is it because it rained(that you are all wet)?”

There are some marginal differences between Korean **-l kes-i-** and Japanese **hazu**, such that the former can marginally be used to ask a deity or scientists, who can control rain-fall. **Hazu** cannot be used in such cases.

5. Extensions or Remaining Issues: Other modals and evidentials

5.1. kamosirenai (possible, epistemic)

• In epistemic use, conditional premises are involved, **kamosirenai** patterns with **hazu-da**

(23) If it rains in Sizuoka, it could rain in Tokyo in one hour (we all know)

It is raining in Sizuoka now

It will rain in Tokyo in one hour-**kamosirenai**

(24) If it rains in Sizuoka, it could rain in Tokyo in one hour (we all know)

It is raining in Sizuoka now

*It could rain in Tokyo in one hour-**kamosirenai**?

√ It will rain in Tokyo in one hour-**kamosirenai-no**?

no must be used so that q can ask the premises

5.2 kamosirenai (possible, non-epistemic)

In non-epistemic use, no conditional premises are involved, thus “**q-kamosirenai?**” asks a simple possibility

In other words, nothing relevant is present in Common Ground for non-epistemic use of **q-kamosirenai?**

Someone who doesn't know anything about the weather today says:

(25) Kyoo-wa ame-ga furu-kamosirenai-desu-ka?
today-TOP rain-NOM fall-possible-COP.HON-Q
“Is there a good possibility that it will rain today?”

6. evidentials

sooda (hearsay)

(26) ame ga huru soodesuka ?

(You hear from someone that) it rains-sooda?

=Did you hear from someone that it is going to rain?

yooda (look as if)

(27) ame ga huru yoodesuka ?

(From the look of the sky) it rains-yooda?

=Is the look of the sky such that it is going to rain?

rasil (hearsay, seems)

(28) ame ga huru rasilidesuka ?

(You hear from someone or the current situation is such) that it rains-rasil?=(Did you hear from someone or the current situation is a typical case such that)

It is going to rain?

-sooda (circumstantial)

(29) ame ga huri soodesuka ?

(From the look of the sky it is likely that) it rains-sooda?

=Is the look of the sky such that it is like to rain?

6. Evidential: Korean -l kes kath-ta

• Pi-ga o-l kes kath-sup-ni-ta.

(From the look of the sky) it is going to rain.

• Pi-ka o-l kes kath-sup-ni-kka?

(Is the look of the sky such that) it is going to rain?

Meta-question about the conversational background (the look of the sky).

7. Summary

non-epistemic modals = no conditional premises are involved

Q → asking simple possibilities

evidential modals = abduction involved

Q → asking presuppositions (Takubo 2009, Saito 2006)

epistemic modals = conditional premises are involved

Q → asking what you already know and leads to contradiction

(i.e. partition for yes/no-Q is not available)

References

- Akatsuka, Noriko. 1985. Conditionals and the Epistemic Scale. *Language* 16.624-639
- Groenendijk, Jeroen and Martin Stokhoff. 1994. Studies on the semantics of questions and the pragmatics of answers. PhD thesis, University of Amsterdam.
- Hamblin, Charles. 1973. Questions in Montague English. *Foundations of Language* 10. Reprinted in: B. Partee (ed.) 1976. *Montague Grammar*. New York: Academic Press.
- Karttunen, Lauri. 1979. Syntax and semantics of questions. *Linguistics and Philosophy* 1, 3-44.
- Kratzer, Angelika. 1981. Notions of Modality. Reprinted with revisions as Ch.2 (pp.27-69) of Kratzer. 2012. *Modals and Conditionals*. Oxford University Press.
- Kratzer, Angelika. 1991. Modality, von Stechow and Wunderlich (eds.) *Semantik: Semantik: Ein internationales Handbuch der zeitgenössischen Forschung* 139-150. De Gruyter.
- Kino, Susumu. 1973. *Structures of Japanese*. MIT Press.
- Oosaka T. 1995. Hontaihaakujasii no setu [Theory of rasi as 'capturing' 'math body']. In: Professors Miyazi Hiroshi and Miyazi Atsuko Sagesei Koki Ninen Ronshuu (Festschrift for Professors Miyazi Hiroshi and Miyazi Atsuko's 70th Birthday). Tokyo: Meiji Syoin: 527-546.
- Saito, Manabu. 2006. *Stengenno no syokoku sutuyoo to tsikkanni* (Evidential reasoning and knowledge management in natural language). Unpublished Ph.D. dissertation, Kyushu University.
- Takubo, Yukinori. 2009. *Conditional modality: Two types of modal auxiliaries in Japanese*. Pizziconi, B. and M. Kizu (eds.) *Japanese Modality: Exploring its Scope and Interpretation*. Palgrave Macmillan.
- Takubo, Yukinori and Summi Kim. 2009. *Kankokugo-to nihongo-no modariti-hyooen-no taishoo* (A comparative study of the modality expressions in Japanese and Korean) *Chosen-hantoo-no kotoba-to shakai* (The Languages and the Society of Korean Peninsula). 285-312. Akashi-shoten.

Psychological Origos and Levels of Deixis

Seungho Nam (Seoul National University)

1. INC and psych predicates: Data

Korean has a constraint on the tense and the subject in psych sentences:

- (1) a. na-nun mucheok cilwuha-e [나는 무척 지루해.]
I-Top very.much bored-Decl
'I am very much bored.'
- b. *Koni-nun mucheok cilwuha-e [*건이는 무척 지루해.]
K-Top very.much bored
'Koni is very much bored.'
- c. Koni-nun mucheok cilwuha-ess-e [건이는 무척 지루했어.]
K-Top very.much bored-Past-Decl
'Koni was very much bored.'
- (2) a. Na-nun ttena-ko.siph-e [나는 떠나고 *싶어]
I-Top leave-BoulM-Decl (BoulM = bouletic modal)
'I would like to leave.'
- b. *Koni-nun ttena-ko.siph-e [*건이는 떠나고 *싶어]
K-Top leave-BoulM-Decl (BoulM = bouletic modal)
'Koni would like to leave.'
- c. Koni-nun ttena-ko.siph-ess-e [건이는 떠나고 싶었어]
K-Top leave-BoulM-Past-Decl
'Koni wanted to leave.'

I-and-Now Condition [INC]:

The subject of psych predicates and bouletic modals should be the first person in the present tense.

Epistemic modals reverses the acceptability of psych sentences:

- (3) a. *na-nun mucheok cilwuha-keyss-e [*나는 (지금) 무척 지루하겠어.]
 I-Top very.much bored-Decl
 ‘I might be very much bored.’
 b. Koni-nun mucheok cilwuha-keyss-e [건이는 무척 지루하겠어.]
 K-Top very.much bored
 ‘Koni might be very much bored.’

Japanese also has a similar constraint (Kuroda 1973)

- (4) a. *Mary-wa sabishii-yo
 M-Top lonely-PRESENT-YO
 ‘Mary is lonely-YO.’ [YO = “reportive style”]
 b. *Mary-wa atukatta-yo
 M-Top hot-Past-YO
 ‘Mary was hot-YO.’
 c. Watashi-wa samui-yo
 I-Top cold-YO
 ‘I am cold.’
 (5) Mary-wa sabishii-no.da
 M-Top lonely-NO.DA
 ‘Mary is lonely.’

→“This constraint is lifted under evidential markers on the clause or the verb and also in nominalized clauses.” (Tenny 2006)

Korean has a rich set of complex psych predicates which consist of a psych noun and a deictic verb - ‘come’ or ‘go’.

[psych noun + ‘come’/‘go’]
 ihay ‘understanding’ + GO = ‘understand’
 coleum ‘drowsiness’ + COME = ‘get sleepy’

- (6) a. Na-nun Jini-ka ihay-ka ka-(ass)-ayo [나는 진이가 이해가 가요/갔어요]
 I-Top J-Nom understanding-Nom go-(Past)-Decl
 ‘I understand/understood Jini.’
 b. Nay-ka coleum-i o-(ass)-ayo [내가 졸음이 와요/왔어요.]
 I-Nom sleepiness-Nom come-(Past)-Decl
 ‘I am/was sleepy.’
 (7)a. Koni-ka/hantey coleum-i o-ass-eyo [건이가 졸음이 왔어요.]
 K-Nom/Dat sleepiness-Nom come-Past-Decl
 ‘Koni was sleepy.’
 b. *Koni-ka coleum-i o-ayo [건이가 졸음이 와요.]
 K-Nom drowsiness-Nom come-Decl
 ‘Koni feels sleepy.’

- c. Koni-ka coleum-i o-keyss-eyo [건이가 졸음이 오겠어요.]
 K-Nom sleepiness-Nom come-EpistMod-Decl
 ‘Koni might feel sleepy.’

Q: What is the nature of the constraint INC?

Q: What determines the use of deictic verbs ‘come’ and ‘go’ in complex psych predicates?

**Q: How can we account for the interaction between psych predicate and epistemic modals?
 Further, what is the general dynamics of modal and deictic perspectives?**

2. Psychological Origos

**There are various types of “perspectives” involved in different levels of sentence structure.
 We are particularly interested in the interaction among the “psychological” perspectives and
 “deictic” referential perspectives.**

- (8) I am afraid he should come to your party.
 - Speaker=Subject: Experiencer perspective, Spatial deictic perspective for come
 - Hearer=[Specifier of party]: Spatial deictic perspective for come

→ Deictic elements (interlocutors) determine the perspectives in different levels of a sentence.

- (A) Evaluator/Judge of Modality
 - Epistemic, Deontic, Bouletic, Teleological, Purposive, etc.
 - Mood and performatives: Dynamics between speaker and addressee
 “Evaluator/initiator” of modal base, ordering source as well as common ground

e.g., rising declaratives: “addressee’s doxastic perspective” Gunlogson (2001, 2002), Farkas & Bruce (2010) -
 presupposing that the addressee is committed to the content of the declarative, according to the common ground.

(B) Evidential origo

Evidentiality needs to be anchored by the evidence-learner/acquirer who sees, hears, infers, reports, and so on.
 The perspective-holder whose experience is purported to be expressed by an evidential

Evidentiality involve cognitive and perceptual experience, and refers to its perception by a conscious subject:

- The “conscious subject/experiencer” has been labeled in various ways in the literature: judge, experiencer
 (Mushin 2001), evidential origo (Garrett 2001), witness (Speas 2004), observer (Aikhenbald 2004a), perceiver (Speas
 2004), etc.

Evidentiality interacts with deixis.

- (i) Evidential origo = speaker
 - Often the perceptual subject is co-identified with the speaker:

- Evidentiality as a deictic category: marking relationships between speakers and events/situations.
(Cf. modality is defined often as denoting a speaker's attitude or belief: epistemic, bouletic, deontic, etc.)

(ii) Evidential origo \neq speaker

- Cross-linguistic research shows many languages use non-speaker's evidentiality.

(C) Attitude holders for the complement clauses

Harry believes that Mary should come to the party.

- believe, know, hope, wish, realize, discover, surprised, etc.

(D) Judge for evaluative expressions

Bill often enjoys a tasty fast food.

- tasty/fun, etc. Farkas and Bruce (2010)

“speakers discourse commitment, i.e., doxastic perspective”

(E) Experiencer of psychological event/state

Children fear thunderstorm. [Experiencer – Stimulus]

The boy was cold outside.

(F) Logophoric reference

- anaphors, pronouns, definite descriptions, deictic expressions, etc.

Mary wants me to come to John's party.

- Sells' (1987) logophoric hierarchy: the referent of a logophoric pronoun can bear one of the three pragmatic roles, which fall into a [typological] hierarchy: source >> self >> pivot

source: the one who makes the report

self: the one whose 'mind' is being reported

pivot: the one from whose physical point of view the report is made

(G) Perspective holder of appositives

speaker oriented vs. non-speaker oriented appositives

[Context: Chris and Ann are talking. Ann says:]

Ron, (who is) a government worker, hates the government.

(Barlew 2017: 296) “non-speaker oriented appositives”

(Amaral et al. 2007, Harris and Potts 2009)

(H) spatial perspectives

- Deictic motion verbs 'come' and 'go':

Mary wants me to come to John's party.

- Deictic locatives (Nam 1995):

Mary ran through the forest (from here). [speaker's perspective]

Come across the street (from there, the bus stop). [addressee's perspective]

Mary is standing in front of/behind the tree (from here/there).

Let us use the term ‘origo’ for the various types of perspectives illustrated above, either psychological or spatial ones.

→ Origo is the origin of vector, from which the direction and the magnitude of the vector originate.
(cf. modal base, ordering source, degree of modal force)

Q: We want to see the interaction among the different origos in a sentence.

Q: How do the psychological and the spatial origos intertwined

3. Origo inheritance: Holes, Filters, and Plugs

Speaker (the matrix spatial origo) may or may not control the origo of embedded modal/psych domain.

(A) [[doxastic origo] under speaker]

- (9) a. Jini-nun Koni-ka o-ass-ta.ko mit-eyo [진이는 건이가 왔다고 믿어요]
J-Top K-Nom COME-Past-Comp believe-Decl (from Jini’s/speaker’s perspective)
‘Jini believes that Koni has COME.’
b. Jini-nun Koni-ka ka-ass-ta.ko mit-eyo [진이는 건이가 갔다고 믿어요]
J-Top K-Nom GO-Past-Comp believe-Decl (from Jini’s/speaker’s perspective)
‘Jini believes that Koni has GONE.’

→The doxastic domain of ‘believe’ is a hole for Speaker’s spatial origo. Thus the speaker may control the spatial origo of the embedded event.

(B) [[experiencer origo] under speaker]

- (10) a. Koni-nun Jini-ka phathi-ey o-un.kes-i kippu-ess-eyo [건이는 진이가 온 것이 기뻐어요.]
K-Top J-Nom party-to COME-Nominal-Nom delighted-Past-Decl
‘Koni was delighted that Jini came to the party.’
(from experiencer’s/*speaker’s perspective)
b. Koni-nun Jini-ka phathi-ey ka-un.kes-i kippu-ess-eyo [건이는 진이가 간 것이 기뻐어요.]
K-Top J-Nom party-to GO-Nominal-Nom delighted-Past-Decl
‘Koni was delighted that Jini went to the party.’
(from experiencer’s/*speaker’s perspective)

→ The experiencer domain of ‘delighted’ is a plug for Speaker’s spatial origo. Thus the spatial deixis of the psych event is not governed by the speaker’s perspective.

[[bouletic origo] under speaker]

- (11) a. Jini-nun Koni-ka ceypal o-ki-lul wenha-eyo
J-Top K-Nom please COME-Nominal-Acc want-Decl
'Jini desperately wants Koni to COME.' (from Jini's/*speaker's perspective)
- b. Jini-nun Koni-ka ceypal ka-ki-lul wenha-eyo
J-Top K-Nom please GO-Nominal-Acc want-Decl
'Jini desperately wants Koni to GO.' (from Jini's/speaker's perspective)

→ bouletic domain of 'want' is a filter for Speaker's spatial origo. Thus the speaker may or may not control the spatial origo of the embedded event.

Psychological/Experiencer Origo and Spatial/Speaker Origo

Three groups of constructions depending on inheritability of the matrix Spatial origo [Top-to-bottom inheritance]

Cf. Karttunen, Lauri (1973) 'Presuppositions of Compound Sentences,'

→ three groups of predicates which block/pass/cancel the presuppositions of the complement clause: plugs, holes, and filters [bottom-up projection]

(A) grammaticized modals are holes of Speaker Origo inheritance [come/go]

à The matrix S-origo controls the space of modal domain - control structures.

-tako.mit 'believe' - doxastic verb

-ko.shiph 'would like' (INC) - bouletic origo

-koca.ha 'planning to' (INC) - purposive origo

(B) lexical modals are filters of S-origo inheritance [*come/go] - plug for COME, hole for GO

à The matrix S-origo selectively controls the space of modal domain - non-control structures.

-ki.wenha 'want' - bouletic desire verb

(C) Psych verbs are plugs of S-origo inheritance [*come/*go]

à The matrix S-origo does not controls the space of modal domain - anti-control structures.

-ese/kes-i kippu 'delighted with/to' (INC) - experiencer origo

-ulkka twulyep 'afraid that' (INC)

(D) Non-modal verbs are holes of S-origo 'think' à no origo-clash

-tako sayngkakha 'think' (no control)

4. INC and deixis in complex psych predicates

I-and-Now Condition [INC]:

The subject of psych predicates and bouletic modals should be the first person in the present tense.

(A) Psych verb constructions

(12) b. *Koni-nun mucheok cilwuha-e [*건이는 무척 지루해.]

K-Top very.much bored

‘Koni is very much bored.’

b. *Koni-ka twuthong-i o-ayo [*건이가 두통이 와요.]

K-Nom headache-Nom come-Decl

‘Koni has headache.’

(B) Bouletic modal context

(13) a. *Koni-nun ttena-ko.siph-e [*건이는 떠나고 *싶어]

K-Top leave-BoulM-Decl (BoulM = bouletic modal)

‘Koni would like to leave.’

b. Koni-nun ttena-ko.siph-ess-e [건이는 떠나고 싶었어]

K-Top leave-BoulM-Past-Decl

‘Koni wanted to leave.’

(C) Purposive modal context

(14) a. *Koni-ka ttena-koca-ha-eyo [*건이가 떠나고자 해]

K-Nom leave-PurpM-Decl (PurpM = Purposive Modal)

‘Koni intends to leave.’

b. Koni-nun ttena-koca.ha-ess-eyo [건이는 떠나고자 했어요]

K-Top leave-PurpM-Past

‘Koni intended to leave.’

c. na-nun ttena-ulla-(**ess*)-eyo [나는 떠날래요/떠나려다]

I-Top leave-PurpM-(**Past*)-Decl

‘I intend to leave.’

(D) Evidential context

l Some psych verbs observe INC under an evidential domain.

e.g., cilwuha- ‘be bored’, chwup- ‘cold’, thongcung-i o- ‘feel pain’

(15) a. nay-ka cilwuha-te-la [내가 지루하더라]

I-Nom bored-Evid-Decl

‘I was bored-EVID.’

b. *nay-ka cilwuha-kess-te-la [*내가 지루하겠더라]

I-Nom bored-Epis-Evid-Decl

‘I might be bored-EVID.’

Some verbs do not observe INC under an evidential domain.

e.g., cip-ey ka- ‘go home’, molu- ‘not know’, al- ‘know’

- (16) a. *nay-ka cip-ey ka-te-la [*내가 집에 가더라]
 I-Nom home-to go-Evid-Decl
 'I was going home-EVID.'
- b. Jini-ka cip-ey ka-te-la [진이가 집에 가더라]
 J-Nom home-to go-Evid-Decl
 'Jini was going home-EVID.'
- (17) a. *nay-ka ku kes-ul molu-te-la[*내가 그것을 모르더라]
 I-Nom that fact-Acc not.know-Evid-Decl
 'I didn't know that fact-EVID.'
- b. Jini-ka ku kes-ul molu-te-la [진이가 그것을 모르더라]
 J-Nom that fact-Acc not.know-Evid-Decl
 'Jini didn't know that fact-EVID.'

Q: What is characteristics of the two classes of predicates?

(A) INC predicates: requires 1stpersonsubjectinpresenttensesentencesunderEvidentialdomain.

cilwuha- 'bored', kothongsuleop- 'painful', chwup- 'cold', colli- 'sleepy', kippu- 'delighted', twulyep- 'fearful', sulphu- 'sorrow', kulip- 'longed for', sangsimha- 'broken-hearted', kominha- 'distressed', nukki- 'feel', kamdongha- 'be touched', kkaytat- 'realize', somangha- 'hope', kyengilop- 'amazed with' etc.

→ The above psych predicates refer to a state which should be perceived by inward perception by Evidential Origo.

(B) Anti-INC predicates: requires 3rdpersonsubjectinpresenttensesentencesunderEvidentialdomain.

al- 'know', ihayha- 'understand', swukungha- 'consent', naptukha- 'be convinced/persuaded', mit- 'believe', uysimha- 'doubt', tongceingha- 'sympathize'; ttwi- 'run', ca- 'sleep'; po- 'see', tut- 'listen'; cap-hi- 'be caught';

→ The above predicates refer to a state/event which may be perceived by outward observation/hearing/touching by Evidential Origo.

→ The above predicates refer to a state/event which may be perceived by internal observation/hearing/touching by Evidential Origo.

Their semantic features have a correlate with their difference in the use of their Psych Nouns.

Epistemic Origo/Domain interacts with (saves or kills) the embedded Experiencer Origo.

- (18) a. *Jini-ka cilwuha-Ø-eyo [*진이가 지루해요.]
 J-Nom bored-Evid-Decl
 'Jini was bored-EVID.'
- b. *Jini-ka cilwuha-te-la [*진이가 지루하더라.]
 J-Nom bored-Evid-Decl
 'Jini was bored-EVID.'

→ (Zero) Evidential requires INC for Exp-Origo.

(18a') *[3Exp-Origo] 1Evid-Origo]

(18b') *[3Exp-Origo] (1Epist-Origo)] 1Evid-Origo]

→ (Zero) Evidential requires INC for Exp-Origo/Boul-Origo/Purp-Origo/Mirative-Origo.

(19) a. Jini-ka cilwuha-kess-te-la [진이가 지루하겠더라]

J-Nom bored-Epis-Evid-Decl

'Jini might be bored-EVID.'

→ Epist-Origo (Speaker) saves non-1stpersonExp-Origo.[INCexempted]

(19a') [3Exp-Origo] 1Epist-Origo] 1Evid-Origo]

(20) a. nay-ka cilwuha-te-la [내가 지루하더라.]

I-Nom bored-Evid-Decl

'I was bored-EVID.'

b. *nay-ka cilwuha-kess-te-la [*내가 지루하겠더라]

I-Nom bored-Epis-Evid-Decl

'I might be bored-EVID.'

'I was bored-EPIS-EVID.'

→Intervening Epist-Origo (Speaker) kills 1stpersonExp-Origo.

(20b') *[1Exp-Origo] 1Epist-Origo] 1Evid-Origo]

Deixis in complex psych predicates

[psych noun + 'come'/'go']

ihay 'understanding' + GO = 'understand'

coleum 'drawsiness' + COME = 'get sleepy'

(21) a. Na-nun Jini-ka ihay-ka ka-(ass)-ayo [나는 진이가 이해가 가요.]

I-Top J-Nom understanding-Nom go-(Past)-Decl

'I understand/understood Jini.'

b. Nay-ka coleum-i o-ayo [내가 졸음이 와요.]

I-Nom sleepiness-Nom come-(Past)-Decl

'I am/was sleepy.'

Q: What determines the use of deictic verbs 'come' and 'go' in complex psych predicates?

Q: How can we account for the interaction between psych predicate and epistemic modals?

Further, what is the general dynamics of modal and deictic perspectives?

(Zero) Evidential requires INC for Exp-Origo.

*[3Exp-Origo] 1Evid-Origo]

*[3Exp-Origo] (1Epist-Origo)] 1Evid-Origo]

(Zero) Evidential requires INC for Exp-Origo/Boul-Origo/Purp-Origo/Mirative-Origo(?).

(A) INC predicates:

→ requires 1stpersonsubjectinpresenttensesentencesunderEvidentialdomain.

→ INC predicates refer to a state which should be perceived by inward perception by Evidential Origo.

→ They derive psych nouns which selects the light verb COME to form a complex psych predicate:

cilwuham ‘boredom’, Kothong ‘pain’, chwuwi ‘coldness’, coleum ‘drowsiness’, kippum ‘delight’, twuleywum ‘fear’, sulphum ‘sorrow’, kuliwum ‘longing’, sangsim ‘heart break’, komin ‘stress’, nukkim ‘feeling’, kamdong ‘touching’, kkaydaleum ‘realization’, somang ‘hope’, miteum ‘belief’, kyengikam ‘amazement’, etc.

(B) Anti-INC predicates:

→ requires 3rdpersonsubjectinpresenttensesentencesunderEvidentialdomain.

→ Anti-INC predicates refer to a state/event which may be perceived by outward observation/hearing/touching by Evidential Origo.

→ They derive psych nouns which selects the light verb GO to form a complex psych predicate:

ihay ‘understanding’, swukung ‘consent’, naptuk ‘convince’, miteum ‘confidence/trust’, uysim ‘doubt’, dongceng ‘sympathy’, etc.

References

- Aikhenvald, A.Y., 2004a. Evidentiality. Oxford University Press, Oxford.
- Amaral, P. and Roberts, C., and Smith, E. A. 2007. Review of The Logic of Conventional Implicatures by Chris Potts. *Linguistics and Philosophy* 30.6: 707-749.
- Barlew, Jefferson. 2017. The semantics and pragmatics of perspectival expressions in English and Bulu: The case of deictic motion verbs. Doctoral dissertation. Dept. of Philosophy, Ohio State University.
- Chung, Kyung-Sook. 2007. Spatial deictic tense and evidentials in Korean. *Natural Language Semantics* (2007) 15:187 - 219.
- Farkas, D. F. & Bruce, K. 2010. On reacting to assertions and polar questions. *Journal of Semantics* 27.1: 81-118.
- Garrett, E.J., 2001. Evidentiality and Assertion in Tibetan (PhD dissertation). University of California, Los Angeles.
- Gunlogson, C. 2001. True to form: Rising and Falling Declaratives as Question in English (PhD dissertation). University of California, Santa Cruz.
- Gunlogson, C. 2002. Declarative questions. *Proceedings of SALT* 12: 124-143.
- Harris, J. A. and Potts, C. 2009. Perspective shifting with appositives and expressives. *Linguistics and Philosophy* 32.6: 523-552.

- Karttunen, Lauri. 1973. Presuppositions of Compound Sentences. *Linguistic Inquiry* 4.2: 169-193.
- Kuroda, S.-Y. 1973. Where Epistemology, Style, and Grammar Meet: A Case Study from Japanese, in Stephen Anderson and Paul Kiparsky (eds.), *A Festschrift for Morris Halle*, Holt, Rinehart and Winston, New York, pp. 377 - 391.
- Lee, Chungmin. 1976. Cases for Psychological Verbs in Korean. *Korean Journal of Linguistics* 1.
- Lee, Chungmin. 1999. Aspects of aspect in Korean psych-predicates: Implications for psych-predicates in general. W. Abraham and L. Kulikov (eds) .) *Tense-Aspect, Transitivity and Causativity*. John Benjamins.
- Mushin, I., 2001. *Evidentiality and Epistemological Stance: Narrative Retelling*. John Benjamins, Amsterdam.
- Nam, S. 1995. *The Semantics of Spatial Prepositional Phrases in English* (PhD dissertation). University of California, Los Angeles.
- San Roque, Lila, Simeon Floyd, and Elisabeth Norcliffe. 2017. Evidentiality and Interrogativity. *Lingua* 186-187: 120-143.
- Sell, Peter. 1987. Aspects of Logophoricity. *Linguistic Inquiry* 18.3: 445-79.
- Speas, Margaret. 2004. Evidentiality, logophoricity and the syntactic representation of pragmatic features. *Lingua* 114: 255-276.
- Tenny, Carol. 2006. Evidentiality, Experiencer, and the Syntax of Sentience in Japanese. *Journal of East Asian Linguistics* 15:245 - 288.

Korean evidential '-te'

and causal relations between situations

Dongsik Lim (Hongik University)

In this presentation, as an extension of Lim (2014, 2018), I argue that -te should be analyzed in terms of situations and relations between them. Specifically, following Kalsang et al. (2013) I argue that the semantics of -te encodes the relation between the situation against which the prejacent is evaluated (the Evaluation Situation) and the situation which consists in the speaker's evidence (the Information Situation), but following Hoe et al. (2018), I also argue that the notion of the Evaluation Situation should be revised in the sense that the Uttering Situation should also be considered, and to give an empirically adequate semantics to Korean evidential -te, the lumping of situations (Kratzer 1986) in the Evaluation Situation as well as in the Information Situation should be done in terms of causal (or weakly, temporally natural and expected) relations between situations. Given this, I try to show that the proposal made here can explain various puzzles regarding -te, including variable evidentiality (Chung 2007, a.o.), temporal interpretation (Lee J. 2013, a.o.), as well as its interaction with lexical aspects (Lim 2014, a.o.).

Subjunctive complementizer in Korean: the interaction of inquisitiveness and nonveridicality

Arum Kang (Korea University) and Suwon Yoon (UT Arlington)

1. Introduction

The goal of current work is to identify the novel type of subjunctive complementizer in Korean. Crosslinguistically, subjunctive mood selection refers to the phenomena whether subjunctive is generally marked in the complement clauses by means of subjunctive marking devices such as overt verbal inflections. As Portner (2018, pp. 78) puts, “mood is a kind of dependent modal element, similar to agreement ... we could say that subjunctive “agrees with” *want* but not *believe*.” For example, as shown below, the volitional verbs *veut* ‘want’ in French and *thelo* ‘want’ in Greek obligatorily select subjunctive whereas the factive verb *know* does not (Giannakidou and Mari 2017, (2b), (5)):

- (1) a. Marc sait que le printemps ***soit/est** arrivé [French]
 Marc knows that the spring be-SBJV-3sg/be-IND-3sg arrived
 ‘March knows that spring has arrived.’
 b. Marc veut que le printemps **soit/*est** long
 Marc wants that the spring be-SBJV-3sg/be-IND-3sg long
 ‘March wants spring to be long.’
- (2) a. **O** Pavlos kseri ***na/oti** efije i Roxani. [Greek]
 The Paul knows-3SG that-SUBJ/that-IND lef-3SG the Roxani
 ‘Paul knows that Roxanne left.’
 b. Thel o **na/*oti** kerdisi o Janis.
 want-1sg that.SBJV/that.IND win.NONPAST-3SG the John
 ‘I want John to win.’

The following lists show indicative/subjunctive governors (Portner 2018, pp. 73-74):

- (3) Indicative governors (mood selection in complement clause):
- a. predicates of knowledge and belief (e.g. *know, believe*)
 - b. predicates of assertion (e.g. *say*)
 - c. **predicates of inquiry (e.g. *ask*)**
 - d. natural factive predicates (e.g. *remember*)
 - e. predicates of fiction and mental creation (e.g. *dream*)
 - f. commissive predicates (e.g. *promise*)
 - g. perception predicates (e.g. *see*)
- (4) Subjective governors (mood selection in complement clause):
- a. **Predicates of inquisitiveness (e.g. *wonder*)**
 - b. Preference predicates (e.g. *want, hope*)
 - c. Directive predicates (e.g. *demand*)
 - d. causative and implicative predicates (e.g. *make*)
 - e. negative counterparts of predicates of knowledge/belief and predicates of assertion (e.g. *doubt*)
 - f. modal predicates (e.g. *necessary, possible, probable*)

The above data provides an important insight on the necessity of its extended spectrum of subjunctive as follows: First, subjunctive mood can be marked not only on the verbs in an embedded clause in (1b) but also on the sentential particles such as the subordinator *C* in (2b). Second, as shown in (3c) and (4a), indicative/subjunctive governors can involve rogative predicates: the predicates of inquiry *ask* is the interrogative-selecting counterparts of verbs of assertion (*ask*: ‘want to be told.’) whereas the predicate of inquisitiveness *wonder* is the interrogative-selecting counterparts of predicates of belief/knowledge (*wonder*: ‘want to know’) (Portner 2018, pp. 72).

Although the study of subjunctive mood has been extensively conducted in a variety of European languages, it has been noted that Korean also employs subjunctive markers. For example, Shim (1995) suggests that the Korean particle *ki* functions as a subjunctive variant of indicative complementizer. Further, Yoon (2011, 2013) has developed the pioneering studies of subjunctive by establishing empirical and theoretical correlations between subjunctive and evaluative negation (EN). As far as our knowledge goes, however, the precise nature of complementizer and its relations with subjunctive mood and attitude predicates have yet to be discussed. In this paper, we mainly focus on subjunctive marking occurring in the interrogative subordinator *C*.

Korean employs two types of overt particles to mark interrogative complementizers. As shown below, we have the ordinary interrogative complementizer *ci* in (5a) and (6a) and modalized interrogative complementizer (*u*)-*l*-*kka* in (5b) and (6b):

- (5) a. Mina-nun Chelswu-ka pathi-ey o-nun-**ci** **kwungkumha**-ass-ta.
M.-Top C.-Nom party-Loc come-Asp-Q. Comp wonder-Past-Decl
‘Mina wondered whether Chelswu would come to the party.’
- b. Mina-nun Chelswu-ka pathi-ey o-**l-kka** **kwungkumha**-ass-ta.
M.-Top C.-Nom party-Loc come-Mod-Q.Comp wonder-Past-Decl
‘Mina wondered if Chelswu **might** come to the party.’
- (6) a. Mina-nun Chelswu-ka pathi-ey o-nun-**ci** **mulepo**-ass-ta.
M.-Top C.-Nom party-Loc come-Asp-Q.Comp ask-Past-Decl
‘Mina asked whether Chelswu would come to the party.’
- b. #Mina-nun Chelswu-ka pathi-ey o-**l-kka** **mulepo**-ass-ta.
M.-Top C.-Nom party-Loc come-Mod-Q.Comp ask-Past-Decl
‘(lit.) Mina asked if Chelswu **might** come to the party.’

As shown above, the difference of *ci* and *(u)l-kka* comes from two aspects: First, unlike *ci*, *(u)l-kka* exhibits a distributional restriction. It freely co-occurs with the inquisitive verb *kwungkumha* ‘wonder’, but not with the inquiry verb *mulepo* ‘ask’. Second, the sentence of *(u)l-kka* in (5b) contains an epistemic modal ‘might’ in the subordinate clause, whereas *ci* in (5a) and (6a) does not exhibit such modal property.

This brings us to the main question of the present work, as follows: First, what is the semantic contribution of *(u)l-kka*? Second, how its behavior is different from the ordinary interrogative *ci*-complementizer? Given these questions, we want to explore the empirical dimension and figure out how the distinct behaviors in Korean and other languages relate to each other. As will be seen in detail soon, we argue that *(u)l-kka* is a grammatical category of sub-type of subjunctive in the interrogative complement clause. In this vein, from now on, we term it as a subjunctive-interrogative complementizer (SUBJ.Q-Comp, henceforth), and gloss it as ‘whether.SUBJ’.

The behavior of *(u)l-kka* SUBJ.Q-Comp is distinct from the cross-linguistic subjunctive in many aspects. For reasons to be made clear soon, we provide three core properties as follows: First, Korean subjunctive mood can appear in the interrogative complement clause. Second, as revealed in (5) and (6), *(u)l-kka* SUBJ.Q-Comp and *ci* Q-Comp does not exhibit complementary distribution of predicates in mood selections. It is contrasted with the strict selection in other languages. Third, as a subjunctive particle, the addition of *(u)l-kka* manifests an epistemic weakening on the proposition. Specially, we will show that *(u)l-kka* exhibits speaker/subject’s epistemic uncertainty in her doxastic space. In this vein, unlike the typical mood selection, *(u)l-kka* SUBJ.Q-Comp appear to add semantic effects, which Giannakidou calls evaluative (Giannakidou 2015). Terming it subjunctive mood choice, we claim that *(u)l-kka* functions as an *epistemic subjunctive*.

The paper proceeds as follows: In Section 2, we discuss the core properties of *(u)l-kka* by examining the contrast between *(u)l-kka* SUBJ.Q-Comp and *ci* Q-Comp. In Section 3, we present the semantic analysis of *(u)l-kka*-complementizer, which is sensitive to epistemic subjunctive mood. We conclude in Section 4 with the theoretical implications.

2. The core properties of *(u)l-kka*

2.1 Predicates that take interrogative complements in English

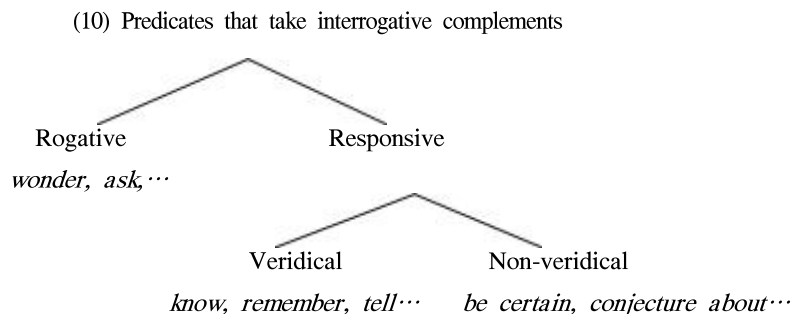
Traditionally, interrogative predicates requiring clausal complements fall in different categories depending on the kind of complements that they can embed. These contrasts in embedding are widely supported by cross-linguistic data (Karttunen 1977; Lahiri 2002). For example, in English, a verb like *believe* in (7) takes only declarative complements while a verb like *wonder* takes interrogative complements in (8):

- (7) a. Mina believes that today is John's birthday.
b. *Mina believes whether today is John's birthday.
- (8) a. *Mina wonders that today is John's birthday.
b. Mina wonders whether today is John's birthday.

Above data reveal that there is a syntactico-semantic dichotomy in assigning different semantic values to questions and assertions. However, not all predicates exhibit such a complementary distribution. For example, as a responsive predicate, the factive verb *know* takes both declarative and interrogative complement:

- (9) a. Mina knows that today is John's birthday.
b. Mina knows whether today is John's birthday.

For the classification of question embedding predicates, we build on the Lahiri (2002)'s category as follows (Lahiri 2002, pp. 286-287):



2.2 The different types of attitude predicates with *ci* and *(u)l-kka*

The behavior of *(u)l-kka* SUBJ.Q-Comp appears to make a crucial distinction from *ci* Q-Comp as follows: First, *(u)l-kka* has a dual use of the modalized question (MQ, henceforth: Kang 2015; Kang and Yoon, to appear) in a non-embedded clause. Unlike regular questions, MQs express speaker's epistemic uncertainty on the propositional content, and ask about the possibility of the propositional content. As shown below, given that we are aware that Santa Clause does not exist in the real world, it is odd to ask a question using an ordinary question form as in (11) because it would mean that the speaker expects an actual answer. On the other hand, *(u)l-kka* in (12) is properly used in a given context:

Context: Mina and Chelswu are talking about Santa Clause. They wonder how old Santa would be if he exists. Mina asks Chelswu:

- (11) *santa-nun myech-sal-i-l-kka?*
 Santa-Top what-age-be-Mod-Q
 'How old **might** Santa Clause be?'
- (12) *#santa-nun myech-sal-i-ni?*
 Santa-Top what-age-be-Q
 'intended: How old is Santa (if he exists?)'

Second, unlike *ci* in (13), *(u)l-kka* cannot combine with the inquiry verb *mulepo* 'ask' in (14):

- (13) *Mina-nun Chelswu-ka pathi-ey o-nun-ci mulepo-ass-ta.*
 M.-Top C.-Nom party-Loc come-Asp-Q.Comp ask-Past-Decl
 'Mina asked whether Chelswu would come to the party.'
- (14) *#Mina-nun Chelswu-ka pathi-ey o-l-kka mulepo-ass-ta.*
 M.-Top C.-Nom party-Loc come-Mod-Q.Comp ask-Past-Decl
 '(lit.) Mina asked if Chelswu might come to the party.'

Third, whereas *ci* co-occurs with the knowledge (i.e. factive) verb *al* 'know' in (15), the occurrence of *(u)l-kka* is strictly restricted with the verb *al* 'know' in (16):

- (15) *Mina-nun Chelswu-ka pathi-ey o-l-ci al-ko.iss.ess-ta.*
 M.-Top C.-Nom party-Loc come-Mod-Q.Comp know-Asp-Decl
 'Mina knew that/whether Chelswu would come to the party.'
- (16) *#Mina-nun Chelswu-ka pathi-ey o-l-kka al-ko.iss.ess-ta.*
 M.-Top C.-Nom party-Loc come-Mod-Q.Comp know-Asp-Decl
 '#(lit.) Mina knew if Chelswu might come to the party.'

Forth, the *(u)l-kka*-clause cannot take the predicate such as *hwaksinha* 'be certain' in (18), which is unexpected from the ordinary interrogative complementizer *ci* in (17):

- (17) Mina-nun Chelswu-ka pathi-ey o-l-**ci** hwaksinha-ss-ta.
 M.-Top C.-Nom party-Loc come-Mod-Q.Comp be.certain-Past-Decl
 ‘Mina was certain whether Chelswu would come to the party.’
- (18) #Mina-nun Chelswu-ka pathi-ey o-l-**kka** hwaksinha-ss-ta.
 M.-Top C.-Nom party-Loc come-Mod-Q.Comp be.certain-Past-Decl
 ‘#(lit.) Mina was certain if Chelswu might come to the party.’

Fifth, the addition of *(u)l-kka* appear to come with a semantic effect. For example, when *(u)l-kka* combines with the polysemous verb *siph* in (19), *(u)l-kka* necessarily *choices* the doxastic related meaning ‘conjecture’ in (19b). By doing this, the meaning change occurs into dubitative reading. It is contrasted with *ci* in (20):

- (19) *siph* (verb)
- a. want
 - b. conjecture**
 - c. fear, worried
 - d. hope
 - e. intend
- (20) Mina-nun Chelswu-ka pathi-ey o-l-**kka**/#-nun-ci siph-ess-ta.
 M.-Top C.-Nom party-Loc come-Mod-Q.Comp/-Asp-Q.Comp conjecture-Past-Decl
 ‘Mina doubted if Chelswu might come to the party.’

The example in (20) is clearly the case of the verb doubt which expresses an attitude of non-belief. *(U)l-kka* manifests properties analogous to what an irrealis assertion denotes.

Sixth, only *(u)l-kka* is compatible with expletive negation, whose occurrence is restricted to the subjunctive mood environment (Yoon 2013; Choi and C. Lee 2017, a.o.). As shown below, in this case, uncertainty reading remains:

- (21) Mina-nun Chelswu-ka pathi-ey oci-anh-**ul-kka**/#-nun-ci siph-ess-ta.
 M.-Top C.-Nom party-Loc come-Neg-Mod-Q.Comp/-Asp-Q.Comp conjecture-Past-Decl

In light of the above data, we might assume that by using *(u)l-kka*, the speaker makes the non-commitment to the truth of the embedded proposition. To put it differently, as a SUBJ.Q-Comp, the addition of *(u)l-kka* in embedded clauses produces a weakening, nonveridicality effect (Giannakidou 1995, et seq.). By doing this, it weakens the degree of certainty about the proposition in embedded question and gives rise to epistemic uncertainty or doubt interpretation. The epistemic SUBJ.Q-Comp questions the subject’s belief and knowledge and expresses her epistemic uncertainty on the possible answer sets.

2.3 *(U)l-kka* and possible answers

One of the important characteristics of responsive predicates is that they express a relation between the holder of an attitude and a proposition which is an answer to the embedded question.

- (22) a. “Mina knows whether the spring will come” is true iff Mina knows p , where p is the true answer to “will the spring come?”
b. “Mina and Bill agree on whether the spring will come” is true iff Mina and Bill agree that p , where p is a **possible answer** to “will the spring come?”

Given the above observations, Lahiri divides responsive predicates into two classes (adapted from Égré and Spector 2007), as follows:

- (23) a. Veridical-responsive: Responsive predicates that express a relation to the actual true answer
b. Nonv-veridical responsive: Responsive predicates that express a relation to a potential answer

Veridical-responsive predicates express a relation between an attitude holder and a proposition. Here the proposition is the actual complete answer to the embedded question. The sentence “Mina knows that Chelswu is coming to the party” entails that Mina has a true belief as to whether Chelswu is coming to the party. On the other hand, non-veridical responsive predicates express a relation between an attitude holder and a proposition that is simply a potential complete answer to the embedded question. The sentence “Mina is certain whether Chelswu is coming to the party” is true even if Mina believes that Chelswu is coming to the party while in fact he isn’t.

As a SUBJ.Q-Comp, the addition of *(u)l-kka* in embedded clauses produces a weakening, nonveridicality effect. By doing this, it specifies the degree of certainty about the proposition in embedded question and gives rise to epistemic uncertainty or doubt interpretation. Let’s see the example below. As mentioned above, the non-veridical responsive predicates taking *(u)l-kka* express a relation between the holder of an attitude and a proposition which is a possible answer to the embedded question. Given that the semantics of embedded questions comprises all potential answers, the employment of an interrogative complementizer introduces both positive and negative cases as equal possibilities of p or $\neg p$. If the speaker chooses *(u)l-kka*, she additionally expresses the subject’s weakest commitment to the possibility of propositional content:

- (24) “Mina conjectures *(u)l-kka*-Comp Chelswu will come to the party” is true iff Mina conjectures that p , where p is a potential answer to “will Chelswu come to the party?” (interrogative) & Mina has epistemic weakening between the possible answer sets (epistemic uncertainty).

Reporting on the consideration of a set of alternative possibilities, *(u)l-kka* expresses subject’s uncertainty in her doxastic space.

- (25) a. Mina-nun Chelswu-ka pathi-ey oci-anh-**ul-kka** ship-ess-ta.
M.-Top C.-Nom party-Loc come-Neg-Mod-Q.Comp conjecture-Past-Decl
‘Mina conjectured if Chelswu **might** come to the party.’
- b. Mina conjectured (*u*)/-*kka*-Comp Chelswu is coming to the party.
--> true if Mina has weaker belief that Chelswu is coming while she has weaker belief that Chelswu is not coming.
- i.e. Mina conjectured that p, where p is a potential answer to *is Chelswu coming to the party?*
& Mina was undecided as to where the actual world is in p or $\neg p$. (epistemic uncertainty)
- (26) “Mina conjectured (*u*)/-*kka*-Comp Chelswu is coming to the party” is true iff Mina conjectured that p, where p is a potential answer to “is Chelswu coming to the party?” (interrogative) & Mina had uncertainty between the possible answer sets (epistemic uncertainty).

Reporting on the consideration of a set of alternative possibilities, (*u*)/-*kka* expresses subject’s uncertainty in her doxastic space. In this vein, it inevitably involves the subjunctive mood. (*u*)/-*kka* reduces speaker’s commitment to the truth of the sentence. We take this to argue that a strong connection exists between the non-veridical reading and the subjunctive mood since both are related to speaker/subject’s non-commitment to the truth of proposition.

Conjecture and doubt trigger a non-homogeneous doxastic space hence subjunctive is predicted. In this vein, (*u*)/-*kka* can be taken as an epistemic subjunctive complementizer. The epistemic SUBJ.Q-Comp questions the subject’s belief and knowledge and expresses her epistemic uncertainty on the possible answer sets. The epistemic SUBJ.Q-Comp (*u*)/-*kka* presupposes that they contain the subject’s partitioned private space.

3. Analysis

Traditionally, subjunctive is deeply tight to the notion of subjectivity, i.e. consideration of spaces of beliefs, doxastic, epistemic, bouletic (Farkas 1992; Giannakidou 1994 et seq.; Villata 2008; Smirnova 2013, a.o.). Building on Giannakidou (1994 et seq.), we treat (non)veridicality as a property of subjective spaces. The subjective spaces are the based on the epistemic state of an individual as follows:

- (27) Epistemic state of an individual anchor *i* (Giannakidou 1999: (45))

An epistemic state $M(i)$ is a set of worlds associated with an individual *i* representing worlds compatible with what *i* knows or believes.

$M(i)$ is the private space of *i*’s thought, belief and knowledge, and it plays a key role in truth assessment. Subjective veridicality is anchored to an individual’s $M(i)$. In unembedded assertions, *i* is the speaker. In embedding, *i* is the speaker or the subject:

(28) Subjective veridicality (Giannakidou and Mari 2017: (25)):

A function F that takes a proposition p as its argument is subjectively veridical with respect to an individual anchor i and an epistemic state $M(i)$ iff: $\forall w[w \in M(i) \rightarrow w: \{w' \mid p(w')\}]$

(29) Subjunctive as an indicator of nonveridicality (Giannakidou 1999):

(i) The subjunctive is an indicator of a nonveridical epistemic state or modal base, and is selected by expressions that are at least subjectively nonveridical.

(ii) Subjunctive sentences indicate epistemic weakening.

The subjunctive thus produces epistemic weakening. Commitment weakening is the creation of a nonveridical (i.e. non-homogeneous) epistemic space.

We argue that the licensing of *nka* is revealed in that it indicates the subjective nonveridicality, as follows:

(30) Epistemic subjunctivity marked in subordinator C = subjective nonveridicality

Aforementioned in (28) and (29), subjective nonveridicality means that i is in a state of uncertainty with respect to p . Some worlds in $M(i)$ support p and some other don't:

(31) Subjective nonveridicality of *(u)l-kka* SUBJ.Q-Comp: A function F that takes a proposition p as its argument is subjectively nonveridical with respect to an individual anchor i an epistemic state $M(i)$ iff: $\exists w' \in M(i): \neg p(w') \wedge \exists w'' \in M(i): p(w'')$

The proposed semantics shows how *nka* expresses the speaker's perspective towards p by achieving a partition in the modal base, characterized as a partitioned epistemic space.

4. Conclusions

In this paper, we identified *(u)l-kka* as a novel type of epistemic subjunctive complementizer. We showed that there are three distinct mechanisms for the occurrence of Korean subjunctive: First, it can be formally marked at the level of interrogative subordinator C . Second, subjunctive marking produces semantic contribution of epistemic weakening more than reflecting modal properties of the context in which it occurs. In the sense of mood choice, we revealed the existence of tendency in which the occurrence of *(u)l-kka* depends on the higher verbs whose subject provides nonhomogeneous doxastic space.

Theoretical implication of the current analysis is as follows: First, Korean reveals that subjunctive complement clause does not form the uniform class. In the sense of Portner (2018), subjunctive mood can occur in the "inquisitive" complement clause. Second, in Korean, there is a semantic dichotomy in assigning different semantic values to the ordinary Q-Comp *ci* and the SUBJ.Q-Comp *(u)l-kka*. The parallel contrast between indicative and subjunctive in other languages can be found in *(u)l-kka* and *ci* in Korean. Much more needs to be said to gain a full understanding of the precise nature of the relationship between inquisitiveness and non-veridical subjectivity.

References

- Choi, Yoonhee & Chungmin Lee (2017), Expletive negation and polarity alternatives, In C. Lee, F. Kiefer and M. Krifka (eds.), *Contrastiveness in Information Structure, Alternatives and Scalar Implicatures*, 175-201, Springer.
- Farkas, Donka (1992), On the semantics of subjunctive complements, In P. Hirschbuhler and K. Koerner (eds.), *Romance Languages and Modern Linguistics Theory*, Amsterdam: John Benjamins, 71-104.
- Giannakidou, Anastasia (1994), The Semantic licensing of NPIs and the Modern Greek subjunctive, in Ale de Boer, Helen de Hoop & Henriette de Swart (eds.), *Language and Cognition 4, Yearbook of the Research Group for Theoretical and Experimental Linguistics*, University of Groningen, 55-68.
- Giannakidou, Anastasia (1995), Subjunctive, habituality and negative polarity items, *Semantics and Linguistic Theory 5*. 94-111.
- Giannakidou, Anastasia (1999), Affective dependencies, *Linguistics and Philosophy 22*, 367-421.
- Giannakidou, Anastasia (2015), The subjunctive as evaluation and nonveridicality, epistemic subjunctive, and factive-as- emotive, in Blaszcak et al. (eds.), *For: Mood, Aspect and Modality: What is a linguistic Category?* University of Chicago Press.
- Giannakidou, Anastasia & Alda Mari (2017), (Non)veridicality and mood selection in complement clauses, in Veridicality in grammar and thought: modality, propositional attitudes and negation, ms. University of Chicago.
- Kang, Arum (2015), *(In)definiteness, disjunction and anti-specificity in Korean: a study in the semantics-pragmatics interface*, University of Chicago Ph.D. dissertation.
- Kang, Arum & Suwon Yoon (to appear), From inquisitive disjunction to nonveridical equilibrium: modalized questions in Korean, *Linguistics*.
- Karttunen Lauri (1977), Syntax and semantics of questions, *Linguistics and Philosophy 1*, 3-44.
- Lahiri, Utpal (2002), *Questions and answers in embedded contexts*, Oxford Studies in Theoretical Linguistics, Oxford.
- Portner, Paul (2018), *Mood*, Oxford Surveys in Semantics and Pragmatics, Oxford.
- Shim, sang-wan (1995), A study on the Korean subjunctive complementizer '-ki' (written in Korean), *Language research 31*(1), 55-85.
- Smirnova, Anastasia (2013), Evidentiality in Bulgarian: Temporality, epistemic modality, and information source, *Journal of Semantics 30*, 479-532.
- Yoon, Suwon (2011), *'Not' in the Mood: the Syntax, Semantics, and Pragmatics of Evaluative Negation*, University of Chicago Ph.D. dissertation.
- Yoon, Suwon (2013), Parametric variation in subordinate evaluative negation: Korean/Japanese versus others, *Journal of East Asian Linguistics 22*, 133-166.

■ 개인 발표 2 ■

발화·인용 구문에 의한

목적 관계 표현

— 유형론적 개관 —

홍재성(서울대학교)

1. 머리말
2. 목적 구문
 - 2.1. 목적 관계 의미론과 목적 구문의 한정
 - 2.2. 목적 구문의 유형론
- 2.3. 목적 개념과 인접 개념영역 / 목적 구문 의미·기능의 연합
 3. 발화·인용 구문
 - 3.1. 발화·인용 구문의 범언어적 다양성
- 3.2. 발화·인용 구문의 문법화 / 의미·기능의 연합
 4. 목적 구문과 발화·인용 구문
 - 4.1. 두 구문 간 관계의 양상
 - 4.2. 발화·인용 구문에 의한 목적 관계의 표현
 - 4.3. 유형론적 기술 모형
 - 4.4. 사례의 몇 경우
5. 연구 프로그램의 구성

- ◇ 자료
- ◇◇ 약자·약호
- ◇◇◇ 언어 자료
- ◇◇◇◇ 참고 문헌

1. 자료

(1) ㄱ. (이 연구의 목적은 + 이 연구는 목적이) 이동목적 구문의 범언어적 다양성을 조사하는 (것이다 +데에 있다)

- ㄴ. 나는 이동목적 구문을 언어유형론적 관점에서 연구하는 것이 목적이다.
- ㄷ. 이 연구는 이동목적 구문을 유형론적 관점으로 기술하는 것을 목적으로 한다.
- ㄹ. 이번에는 목적하시는 바를 꼭 이루시기 바랍니다.

(2) ㄱ. 기영이는 책을 빌리-러 도서관에 갔다.

- ㄴ. 지수는 이 문제를 살펴보-(려고+고자+*러) 많은 자료를 수집했다.
- ㄷ. 나는 1차에 합격하-(려고+도록+계) 밤새워 가며 철저히 시험준비를 했다.
- ㄹ. 민기는 동생이 편히 지내-(도록+계) (하기 위해+E) 근처에 혼자서 방을 하나 따로 구했다.
- ㅁ. 민기는 교사가 되-기 위해 교육대학에서 공부했다.
- ㅂ. 기영이는 과제물을 작성하-(느라고+려고+*러) 도서관에서 책을 빌렸다.
- ㅅ. 지수는 집 나간 초코를 찾(아서+으려고+으러) 인근 지역을 모두 돌아다녔다.

(3) ㄱ. 민기는 친구들을 만나-(ㄴ다+겠다+%르거라+%쓰다)-(고+*E) 나갔다.

- ㄴ. 나는 책 사보-(라고+도록+계) 동생에게 용돈을 좀 주었다.
- ㄷ. 나는 피가 나지 (말-라고+않-도록+않-계) 상처를 꼭 처매 주었다.
- ㄹ. 빨리 마르-(라고+도록+계) 양말하고 속옷은 아랫목에 퍼놓았다.
- ㅁ. (기영이는 자기도+나는 나도) 좀 같이 보-(자고+려고) 끼어 들었다.
- ㅂ. 너 편하-(라고+도록+계) 내가 미리 다 손을 써 놓았다.
- ㅅ. 태우는 자기 혼자만 편하-(자고+려고) 그런 잔피를 부린 것이었다.

(4) ㄱ. 민기는 나에게 (기영이를 꼭 만나보라+같이 가자+거기에 가느냐+언제 올 거냐)-(고+E) 했다.

- ㄴ. 민기는 말하기를 자기는 그 모임에 간 적이 없다-(고+E) 했다.

(5) ㄱ. (기영이+너)-도 지수가 좋-(*다+*아+구나+은 거야).

- ㄴ. 그러니까, (기영이+너)-도 지수-(가+를) 보고 싶-(*다+*어+구나+은 거야+다, 이거지?)
- ㄷ. (기영이+*나)-는 지수를 보고 싶어 한다.
- ㄹ. 애는 목이 말라. 빨리 마실 걸 좀 줘라.
- ㅁ. 애는 지수가 좋아. 네가 빨리 좀 만나게 해 주렴.

(6) ABKHAZ

[warà wažə̀ a- gàra (Ø-) q'a- s- c'a- wà- yt' h'a]
you.M now ART-craddle (it-) PREV- ISG- do- DYN-FIN(PRES) QUOTE/PURP

a- žəy-ra w- à- la- ga- wa- yt'
ART-smith-ABST you-it-PREV-begin-DYN-FIN(PRES)

너는 이제 요람을 만들기 위해 장인으로서 작업을 시작한다.

(Hewitt 1987:39)

(7) AGUARUNA

iwi- ya- hi [tipisa -ti tu- sa]
raise.hand-REM.PAST- 1PL.DECL lie.down.PERV-JUSS say-SBD.1PL.SS

우리는 트럭이 멈춰 서-(라고+도록) 손을 들었다.

(Overall 2009:186)

(8) BABUNGO

ŋwə́ jwì féenə̀ [lāa yì gwía yítée nyə̀]
he come.PF here that 3SG.LOG cut.IMPV palmtree DEM

그는 이 야자나무를 자르-(려-고+다-고) 이리 왔다.

(Schaub 1985:39)

(9) CUBEO

ji- RE xiejo-ki- RE epe- Ibā xi- bi'ki-wA bue- I- Rō- I
ISG OBJ child- MS- OBJ put- N/H:3PL ISG.POSS-adult-PL teach.study- STV IN.SG.NMNL-LOC

[xi bāxi- kije-pe a- Iwt ape- Rō xi bāxi- bE- Rō- RE]
ISG.POSS know-PURP.DS say-NFUL.PL.NMNL other-IN.SG.NMNL ISG.POSS-know-NEG- IN.SG.NMNL-OBJ

모르던 것을 배우-(라고+도록) 부모님은 어려서부터 나를 학교에 보내셨다.

(Morse and Maxwell 1999:175)

(10) GAHUKU

Ganaq [kopiɔ zuhaq itune loko] gihisi gomoq nalume
first coffe plant we.FUT saying grass cleaning make.1PL

우리는 커피를 재배하려고 우선 잡초를 뽑는다.

(Deibler 1971:105)

(11) KOBON

Yad ñin rib-em [hagape au- ag- a] a g- em]
1SG hand cut-SS.1SG blood come-NEG-IMP.3SG QUOT do- SS.1SG

ñimagö wam pidöŋ gi- pin
hand bind strong do-PRF.1SG

나는 손을 베어서 피가 나지 말라고 손을 꼭 치매었다.

(Davies 1981:38)

(12) KOMBAL

Nu me- la- ra [ai galemo-fo- nera wa- me- de]
1SG come-stand-SS pig buy- 1PL.INT-QUOT COMPL-come-1SG.NFUT

나는 돼지를 사러 왔다.

(de Vries 1993:96)

(13) USAN

ᵐ. munon [mâni gun-oub (qâmb)] man is- umirei
man yam dig-PL.FUT.SS say.SS garden descend- 3PL.FPST

그 사람들은 암을 캐려고 밭에 내려갔다.

(Reesink 1987:257)

ᵐ. wuri [ai sifeni-on qâmb] ni eng bo mor bif ig- urei
3PL ground dry- SG.IMP say.SS thing DEF again home put.SS COP-3PL.FPST

그들은 땅이 마르라고 이 물건을 집에다 두곤 했다.

(Reesink 1987:258)

(14) EWE

me- tsó ga nê bé(ná) wo- á- ple agbalē
1SG-take money give.3SG PURP 3SG- SUBJ-buy book

나는 그에게 책을 사라고 돈을 (집어) 주었다.

(Heine and Kutera 2002:269)

◇◇ 약자·약호

ABST	abstract
ART	article
ATT	attenuative
COMPL	completive
DEC	declarative
DS	different subject
DYN	dynamic
FIN	finite
FPST	far past
IMP	imperative
IMPF	imperfective
IN	inanimate
INT	intentional
JUSS	jussive
LOC	locative
LOG	logohoric
M	masculine
NEG	negation
NFUT	not-future
N/H	non recent past/present habitual
NMNL	nominalizer
OBJ	object
PL	plural
PERF	perfective
POSS	possesive
PREV	preverb
PURP	purposive
QUOT	quotative
SDB	non-temporal subordinate verb marker
SG	singular
STV	stativizer
SUBJ	subjective

◇◇◇ 언어 자료

Papua New Guinea

Gahuku
Golan
Hatam
Kewa
Kobon
Kombai
Korafe
Manambu
Maybrat
Siroi
Selepet
Tauya
Telefol
Usan
Uri
Wambon
Yimas

South Asia

Kannada
Marathi
Meithai
Punjabi

Caucasus

Abkhaz
Georgian
Lezgian

Africa

Babungo
Baka
Ewe
Fongbe
Kana
Koranko
Lele
Lingala
Maale
Oromo

South America

Aguaruna
Barasano
Baure
Cubeo
Emerillon
Kwaza
Tariana
Wari'

North/meso. America

Chalcatongo mixtec
Jacalteco
Jamul Tilipay
Maricopa
Tzutujil

Eurasia

Korean
Rumanian
Uzbek

Austronesia

Boumaa Fijian
Fehan Tetun
Tukang Besi

참고 문헌

- 고성환. 2003. 『국어 명령문에 대한 연구』. 역락.
- 권재일. 1998. 『한국어 문법』. 박이정.
- 박만규. 1996. “한국어 문장보어 완형문의 통사-의미적 분석”. 『언어』 21-3. 769-791.
- 이연설. 2016. 『한국어 발화동사 연구』. 역락.
- 채숙희. 2011. “목적의 ‘겠다고’에 대하여”. 『정신문화 연구』 34-3. 205-227.
- 홍재성. 2018ㄱ. 한국어 목적 구문의 유형론적 문법기술. 서울대학교 언어연구소 특강자료.
- 홍재성. 2018ㄴ. 목적 관계 개념의 의미유형론을 위하여. 한국언어유형론학회·한국어 의미학회 공동 학술발표회 자료.
- 홍재성. 2018ㄷ. 이동목적구문 유형론. 학술원 논문집 제57집1호. 대한민국 학술원. 1-48.
- Aikhenvald, Alexandra Y. 2003. *A Grammar of Tariana, from Northwest Amazonia*. Cambridge University Press.
- Aikhenvald, Alexandra Y. 2010. *Imperatives and Commands*. Oxford University Press.
- Coulmas, Florian. ed. 1985. *Direct and Indirect Speech*. Walter de Gruyter.
- Davies, H. John. 1981b. *Kobon*. North-Holland.
- Dayley, Jon P. 1985. *Tzutujil Grammar*. University of California Press.
- Deibler, Ellis. 1971. Uses of the verb ‘to say’ in Gahuku. *Kivung* 4(2). 101-110.
- Dixon, R. M. W. 1988. *A Grammar of Boumaa Fijian*. University of Chicago Press.
- Dol, Philmena H. 1999. *A grammar of Maybrat, a language of Bird’s Head, Irian Jaya, Indonesia*, Ph.D. dissertation, University of Leiden.
- Everett, Daniel L. and Kern, Barbara. 1997. *Wari’: The Pacaas Novos Language of Western Brazil*. Routledge.
- Floyd, Simeon. Elisabeth Norcliffe and Lila San Roque. eds. 2018. *Egophoricity*. John Benjamins.
- Foley, William. 1986. *The Papuan Languages of New Guinea*. Cambridge University Press.
- Foley, William. 1991. *The Yimas Language of New Guinea*. Stanford University Press.
- Frajzyngier, Zygmunt. 1991. The de dicto domain in Language. in: *Approaches to Grammaticalization: Volume I*. Elizabeth Closs Traugott and Bernd Heine Eds. 219-251. John Benjamins.
- Gordon, Lynn. 1986. *Maricopa Morphology and Syntax*. University of California Press.
- Güldemann, Tom and Manfred von Roncador eds. 2002. *Reported Discourse: A meeting ground for different linguistic domains*. John Benjamins.
- Harkins, Jean. 1995. *Desire in language and thought*. Ph.D. dissertation, The Australian National University.
- Haspelmath, Martin. 1993. *A Grammar of Lezgian*. Walter de Gruyter.
- Heine, Bernd and Tania Kuteva. 2002. *World Lexicon of Grammaticalization*. Cambridge University Press.
- Hewitt, Brian. 1987. *The Typology of Subordination in Georgian and Abkhaz*. Mouton de Gruyter.
- Hewitt, Brian George. 1989. *Abkhaz*. Routledge.
- Hewitt, Brian George. 1995. *Georgian: A Structural Reference Grammar*. John Benjamins.
- Lowe, Ivan and Ruth Hurlimann. 2002. Direct and indirect speech in Cerma narrative. In *Güldemann and von Roncador*. 71-90.
- Kung, Susan Smythe. 2007. *A Descriptive Grammar of Huehuetla Tepehua*. PhD Dissertation, University of Texas

at Austin.

- MacDonald, Lorna. 1990. *A Grammar of Tauya*. Walter de Gruyter.
- Morse, Nanay and Michael Maxwell. 1999. *Cubeo Grammar*. Summer Institute of Linguistics.
- Miller, Amy. 2001. *A Grammar of Jamul Tipay*. Walter de Gruyter.
- Overall, Simon E. 2007. *A Grammar of Aguaruna*. Ph.D. dissertation, La Trobe University, Australia.
- Overall, Simon (2009) *The Semantics of Clause Linking in Aguaruna*. In: *The Semantics of Clause Linking: a cross-linguistic typology*. Dixon, R and Aikhenvald, Alexandra Y. Eds. 167-192. Oxford University Press.
- Reesink, Ger. 1987. *Structures and their functions in Usan*. John Benjamins.
- Reesink, Ger P. 1999. *A Grammar of Hatam: Bird's Head Peninsula, Irian Jaya*. Australian National University.
- Rose, Françoise. 2003. *Morphosyntaxe de l'Emerillon, langue tupi-guarani de Guyane française*. Ph.D. dissertation, Université Lumière Lyon 2.
- Schmidtke-Bode, Karsten. 2009. *A Typology of Purpose Clauses*. John Benjamins.
- Sridhar, Shikaripur N. 1990. *Kannada*. Routledge.
- Stirling, Lesley. 1993. *Switch-Reference and Discourse Representation*. Cambridge University Press.
- Van Klinken, Catharina Lumien and Williams-van Klinken, Catharina. 1999. *A Grammar of the Fehan Dialect of Tetun: An Austronesian Language of West Timor*. Australian National University.
- Verstraete, Jean-Christophe. 2008. *The status of purpose, reason and intended endpoint in the typology of complex sentences: implications for layered models of clause structure*. *Linguistics* 46-4: 757-788.
- de Vries, Lourens. 1989. *Studies in Wambon and Kombai: Aspects of Two Papuan Languages of Irian Jaya*. University of Amsterdam.
- de Vries, Lourens. 1993. *Forms and Functions in Kombai, an Awyu Language of Irian Jaya*. Australian National University.
- de Vries, Lourens. 2017. *Greater Awyu Languages of West Papua in Typological Perspective*. In: *The Cambridge Handbook of Linguistic Typology*. Aikhenvald, A. and Dixon, R. 942-966. Cambridge University Press.

Yonsei-ScriptDialog: A Richly Annotated Conversation Corpus in Korean for Dialogue System Development

Choi Ji-Myoung (최지명), Kim HanSaem (김한샘)

Yonsei University

Overview

- Overview of dialogue systems
- Current situation
- Datasets for dialogue system
- *Yonsei-ScriptDialog*: a new dataset in Korean for training dialogue systems
- Limits

Acknowledgement

This work was supported by the National Research Foundation of Korea Grant funded by the Korean Government. (NRF-2017M3CA7A0A7088186)

Current dialogue systems: implementation of pre-defined functions

- Virtual assistant dialogue systems
- Embedded devices with dialogue capabilities (home assistant)



Overview of dialogue systems

- Typology of dialogue systems: for what purpose
 - Task-oriented (or domain-specific) dialogue systems
 - Usually called "dialogue systems"
 - general-purpose (or open-domain) dialogue systems
 - Usually called "(social) chatbot system"
 - Their appeal lies not only in their ability to respond to users' diverse requests, but also in being able to establish an emotional connection with users.
 - Can be used for specific purposes
 - *Yonsei-ScriptDialog* corpus as a training and evaluation data for the latter type of systems

- Typology of dialogue systems: implementation techniques
 - Rule-based systems (e.g. ELIZA 60s, PARRY 70s)
 - Rules and small amount of example sentences
- Corpus-based systems
 - IR-based systems (e.g. MS XiaoIce, Cleverbot)
 - Calculation of utterance similarity (utterance context, user information → retrieving the most similar conversations in a corpus)
 - Utterance and context embeddings (hybrid model)
 - End-to-end learning systems (deep neural nets)
 - Generative model
 - Train data: movie dialogue databases, twitter conversations


ELIZA

Men are all alike.
 IN MOST WAYS
 They're always bugging us about something or other.
 CAN YOU THINK OF A SPECIFIC EXAMPLE
 Well my boyfriend made me come here.
 YOUR BOYFRIEND MADE YOU COME HERE
 He says I'm depressed much of the time.
 I AM SORRY TO HEAR YOU ARE DEPRESSED

```

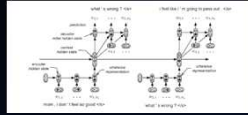
function ELIZA_GENERATOR(user utterance) returns response
  Find the word w in utterance that has the highest keyword rank
  IF w = "you"
    Choose the highest ranked rule r for w that matches utterance
    response ← Apply the transform in r to utterance
  IF w = "my"
    generate ← Apply a transformation from the "history" table list to utterance
    Push generate onto memory stack
  ELSE IF (w is keyword) apply it
  ELSE
    response ← Apply the transform for the "NONE" keyword to utterance
  IF response ← Pop the top response from the memory stack
  return(response)
    
```

XiaoIce

Jerry: "Finding Mr. Right" is a great movie. What do you think XiaoIce?
 XiaoIce: Oh, romantic comedies are my favorite! It stars Tang Wei, who looks like in Seattle. You should watch it.
 Jerry: Do you like Tang Wei?
 XiaoIce: 
 XiaoIce: I guess she stole my heart... and then went off and married
 Tom? to who?
 XiaoIce: Haven't you heard?
 XiaoIce: Tang Wei is engaged to famous Chinese director Kim Tae-hyung

End-to-end system

Sordani et al. "A neural network approach to context-sensitive generation of conversational responses".

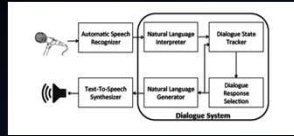


Message (m)	Generated Response (r)
1. I need help. Can you help me?	Of course, I'll be happy to help you. What do you need?
2. I like this book. Do you like it?	That's great! I have the book but I'm not sure I like it. What's the book about?
3. I know! What have you been up to?	Nothing really, just enjoying my summer and working on my thesis.
4. I think you're better than me.	I know, I don't mind you. I love you!
5. Thanks, you're a good friend. Goodbye!	Bye-bye! I hope we see each other again!
6. I agree, but never see him.	That's one of the best, check him out sometime.
7. Today I will walk for miles. Clear sky, clear mind.	I love walking, it's so fun to walk this much. I have a great book, read that!
8. Meeting people over "early events" please. I'm not understanding what the hell did her meeting him is a real pity.	It appears he may have a secret death threat on others while all this was kicking off too. I don't know what's happening to you.

Serban et al. 2015. "Building End-To-End Dialogue Systems Using Generative Hierarchical Neural Network Models".

Workflow of dialogue system

- (Here we are working on the datasets for training the system)



Dialogue system diagram (Serban et al. 2015)

Current situation

- The mainstream is the data-driven (corpus-based end-to-end) systems.
 - Data-driven systems require datasets large enough to train and evaluate.
- Problems
 - Lack of that data
 - Integrating task-oriented and chatbot-style dialogue systems
- That's why we are developing the *Yonsei-ScriptDialog* corpus!

Lack of data

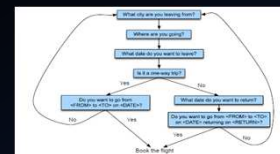
- Gathering large amount of spoken data is expensive, and too time-consuming.
- Privacy issues
- Noisy – can be too discontinuous and fragmentary to be used as input data
- Not so easy to get diverse as required (demographic, situation)

Integration of task-oriented and chatbot-style systems

- Task-oriented system → In-domain conversation
 - Designed to perform specific tasks (next slide)
 - Relatively small data required
 - Lack of flexibility when out-of-domain interactions occur → two strategies ("I have no idea" answers or vague answers)
 - inappropriate answers → frequent conversation interruptions (not trained before) due to unexpected input utterance
- Chatbot style system
 - Diverse human interactions will help deal with out-of-domain interactions.
 - Could improve flexibility through learning from diverse and large conversations

Frame-based dialogue systems

Most commercial digital assistants



Slot	Question
ORIGIN CITY	"From what city are you leaving?"
DESTINATION CITY	"Where are you going?"
DEPARTURE TIME	"When would you like to leave?"
ARRIVAL TIME	"When do you want to arrive?"

Characteristics of the *Yonsei-ScriptDialog* corpus

- Large conversation corpus
 - Resemblance with real dialogue (pseudo-dialogue)
 - "movie language can be regarded as a potential source for teaching and learning spoken language features" (Forchini, 2009)
- Less noisy than real spoken data
- Composed of many conversation situations where only two persons participate (also original multi-party conversations)
- Richly annotated: meta-information, dialogue act, emotions
- Scalable and easy management (easily updatable)

Large conversation data

- Current corpus size (unannotated structured corpus)

No. of lines	No. of eojjeols (words)	No. of utterances	No. of eojjeols from utterances	Mean length of utterance (eojjeol)
356,960	2,587,084	241,960	1,592,759	6.74

- It is increasing now.

Less noisy

- Complex procedures of spoken corpus construction
 - Recruiting → Recording → Sound editing → Transcription → Text editing
 - High potential of noise in each step
 - E.g. quality of recording, recruiting
- Real dialogues are
 - often incomplete, fragmentary and too overlapped
 - Replete with lots of non-standard and slangs
- Script conversations are similar to real dialogues but less noisy.
- Commercial systems use lots of scripted or simulated dialogues.

Bipartite conversation (between two persons)

- Useful for training a dialogue system
- Bipartite conversations in our dataset
 - Dialogue events where only two persons participate
 - Dialogue events between two by transforming the utterance structures of multi-party conversations
 - Tri-turn dialogues: A-B-A type (e.g. A-B-A, A-B-A-B, A-A-B-A, A-B-B-A, ...)

Multi-party conversation

```

[00:00] A: 안녕하세요.
[00:01] B: 안녕하세요.
[00:02] A: 오늘 날씨가 정말 좋네요.
[00:03] B: 맞아요, 정말 좋습니다.
[00:04] A: 어디로 가세요?
[00:05] B: 회사로 가요.
[00:06] A: 회사에서 뭐하나요?
[00:07] B: 고객 상담이죠.
[00:08] A: 고객 상담이 재미있을 거 같아요.
[00:09] B: 아니요, 좀 지루해요.
[00:10] A: 그럼 주말엔 뭐하나요?
[00:11] B: 주말엔 그냥 쉬어요.
[00:12] A: 쉬는 게 제일 좋죠.
[00:13] B: 네, 맞아요.
[00:14] A: 그럼 안녕히 계세요.
[00:15] B: 안녕히 계세요.
                    
```

Bipartite conversation

```

[00:00] A: 안녕하세요.
[00:01] B: 안녕하세요.
[00:02] A: 오늘 날씨가 정말 좋네요.
[00:03] B: 맞아요, 정말 좋습니다.
[00:04] A: 어디로 가세요?
[00:05] B: 회사로 가요.
[00:06] A: 회사에서 뭐하나요?
[00:07] B: 고객 상담이죠.
[00:08] A: 고객 상담이 재미있을 거 같아요.
[00:09] B: 아니요, 좀 지루해요.
[00:10] A: 그럼 주말엔 뭐하나요?
[00:11] B: 주말엔 그냥 쉬어요.
[00:12] A: 쉬는 게 제일 좋죠.
[00:13] B: 네, 맞아요.
[00:14] A: 그럼 안녕히 계세요.
[00:15] B: 안녕히 계세요.
                    
```

```

HT      남산은 왜?
NH      무슨 소리야. 저가 문자 보내놓고.
HT      뭐야?
NH      뭐야 그럼. 너도 똑같은 문자 받았어?
HT      스피린이가 별 거 아니면 끊어. 나 비벼.
NH      비벼봤자. 도상 모퉁에서 여차나 코치고 있었지.
HT      잘 안는거부다? 서교원네?
NH      HT야. 나 그냥 데머의 그 남자와 확 결혼해 버릴까?
HT      경사났네.
NH      나 나름 월영에 살았다고 생각해.
HT      아무런 도감게 떠올거지?
NH      왜 내 이십대의 마지막 하루가 이렇게 끝나야 하니 아!!!!!!
HT      왜!!!
NH      기억났다! 남산! 몇 년 전에 예약 문자 보내왔잖어.
HT      NH
NH      스피린이 마지막 날 남산타워에서 만나자고.
HT      .....아~ 그랬다. 그랬지 참... 하하. 아~ 벌써 그날이 왔네.
NH      그 약속 지켜라!
HT      뭐????
                    
```

Rich annotation

- Meta information (per dialogue event or scene)
 - Demographic information of speakers
 - Gender, age, (occupation), dialect
 - Not all the speakers, but main speakers (lack of reference data)
 - Place: place information where a dialogue occurs
 - Purpose (task): common purpose / task between speakers
 - Topic: dialogue content
- Utterance annotation: DAs (dialogue acts)
 - Utterance intent, function of utterance in a dialogue flow

Purpose / task

- 소개하기 To greet
- 인사하기 To greet
- 상담하기 To consult and be consulted / To give and get consulting
- 신청하기 To apply for public service
- 문의하기 To ask for information and be informed
- 알리기 To give and be given direction
- 안내하기 To invite and be invited
- 초대하기 To invite and be invited
- 약속하기 To make an appointment
- 구매하기 To buy and sell
- 주문하기 To make an order and accept it
- 예약하기 To reserve and accept
- 인터뷰하기 To interview and be interviewed
- 재단하기 To explain and be explained
- 토론하기 To discuss
- 설득하기 To persuade and be persuaded
- 부탁하기 To ask for something

Place

- Airport 공항
- Bus terminal 버스 터미널
- Train station 기차역
- Cafe/Bar 카페, 바
- Home 집
- Hospital/Pharmacy 병원, 약국
- Office 사무실
- Public service center 은행, 관공서 등
- Restaurant 식당
- School 학교
- Street 길
- Shop 상점
- Hotel (lobby) 숙박(호텔 등의 로비, 피난처(데스크))
- Others 기타
- Unknown

DA (dialogue acts)

- Dialogue act annotation is about indicating the kind of intention that the speaker had; what was he trying to achieve?
- Related concepts
 - Domains (= utterance situations)
 - Activity types (cooperative negotiation, information extraction, problem solving, teaching/instruction, counselling, chatting, ...)
 - Sentence types (interrogative, imperative, declarative)
 - Speech acts (← illocutionary force)
- There is no Korean dialogue corpus that is DA-tagged for general-purpose systems.

DA tagset type

- One-dimensional tagging (one tag per utterance) e.g. SWBD-DAMSL
- Multi-dimensional tagging (one or more tags per utterance) e.g. DIT++
 - general-purpose functions + dimension-specific functions (feedback, turn management, time management, ...)
- Units to be tagged
 - Utterances vs. sentences
 - What is an utterance in Korean? How to separate individual utterances from interaction flow (e.g. fragmented parts of an utterance)?

DA tagsets available

- DAMSL (Dialog Act Mark-up in Several Layers): 220 tags
- SWBD-DAMSL - "Application" of DAMSL to the Switchboard data: 42 tags
- ICSI-MRDA: about 60 tags
- MALTUS
- AMI (Decision tree type through yes/no dichotomy): 15 tags
- DIT++ (Dynamic interpretation theory): about 95 tags (42+ general-purpose, 51 domain-specific tags)
- ISO: based on DIT++

Decision on a annotation scheme

- Traum, 2000
 - (1) Theory: the tagset should be related to a theory of the 'functions' that it annotates.
 - (2) Insights from the data: the tagset should be compatible with observations on actual utterances, in a given domain.
 - (3) Empirical validation: the DA set should be reliably tagged by human annotators (high inter-annotator agreement, e.g. using kappa).
 - (4) Possibility of automatic tagging using the tagset, at a reasonable performance level.
 - (5) Role of the application: the tagset should be designed depending on the targeted NLP application (mark relevant 'functions' instead of all 'functions').
 - (6) Mapping to existing DA sets: the DA set should be reasonably compatible with previous tagsets (or at least compared to them) so that useful insights are preserved, and data can be reused.
- In addition to these, two important principles.
 - Not too complex to do annotation on real data
 - Convenient addition or removal tags without affecting the entire scheme
- → We are using the SWBD-DAMSL tagset.

id	turn	act	text	category
1	Statement-completion	sd	Hi, I'm in the legal department.	
2	Acknowledgment	h	Oh, huh.	
3	Statement-completion	sd	I think it's great.	
4	Agreement	aa	That's exactly it.	
5	Appreciation	ba	Yes.	
6	Appreciation	ba	I can imagine.	
7	Yes-No-Question	yn	Do you have to have any special training?	
8	Yes-Answer	ya	Lengthy. I think I learned.	
9	Yes-Answer	ya	Yes.	
10	Conventional-closing	cc	Well, it's been nice talking to you.	
11	Understandable	u	Well, uh, yeah.	
12	Wh-Question	wh	Well, how did you do?	
13	No-Answer	na	No.	
14	Response	rs	Oh, why?	
15	Acknowledgment	ba	No, I don't know if I'm making any sense or not.	
16	Bridge	br	So you can afford to get a house?	
17	Declarative-Yes-No-Question	dn	Well, give me a break, you know.	
18	Declarative-Yes-No-Question	dn	Is that right?	
19	Declarative-Yes-No-Question	dn	You can't be pregnant and have cats.	
20	Summarize/Reformulate	sr	Oh, you mean you switched schools for the kids.	
21	Response	rs	Yes.	
22	Answer	an	Why don't you go first.	

DA annotation procedure

- Semi-automatic procedure: automatic annotation first, then manual correction
 - Very large dataset → nearly unfeasible to conduct manual tagging from scratch
 - Zipf's law in DAs (top 10 DAs out of 42 account for 82% of speech, and top 20, 90%)
- DA prediction system (under experiment)
 - Deep neural network model-based annotator
 - CNN model, CNN + LSTM model
 - No training data in Korean → machine translation as a bridge
 - Trained on SWBD data and evaluated (accuracy: 71.2%, Kappa: 63%) → More diverse data and optimization needed

DA annotation sample test on utterances in Korean

- Experimental test on Korean utterances
 - About 64% (CNN+LSTM) ~ 70% (CNN) success
 - Some bias influence → regularization or uncorrelated features needed or more training data
 - Error type detection (← confusion matrix) → focused correction e.g. *q* → *s*, *ad* → *s*
 - How to split "utterances" in a single turn? (e.g. utterance fragments)

Test sample

utterance	DA	DA Prediction	DA Confusion	DA Precision
1. 안녕하세요. 반갑습니다.	sa	sa	sa	100%
2. 안녕하세요. 반갑습니다.	sa	sa	sa	100%
3. 안녕하세요. 반갑습니다.	sa	sa	sa	100%
4. 안녕하세요. 반갑습니다.	sa	sa	sa	100%
5. 안녕하세요. 반갑습니다.	sa	sa	sa	100%
6. 안녕하세요. 반갑습니다.	sa	sa	sa	100%
7. 안녕하세요. 반갑습니다.	sa	sa	sa	100%
8. 안녕하세요. 반갑습니다.	sa	sa	sa	100%
9. 안녕하세요. 반갑습니다.	sa	sa	sa	100%
10. 안녕하세요. 반갑습니다.	sa	sa	sa	100%
11. 안녕하세요. 반갑습니다.	sa	sa	sa	100%
12. 안녕하세요. 반갑습니다.	sa	sa	sa	100%
13. 안녕하세요. 반갑습니다.	sa	sa	sa	100%
14. 안녕하세요. 반갑습니다.	sa	sa	sa	100%
15. 안녕하세요. 반갑습니다.	sa	sa	sa	100%
16. 안녕하세요. 반갑습니다.	sa	sa	sa	100%
17. 안녕하세요. 반갑습니다.	sa	sa	sa	100%
18. 안녕하세요. 반갑습니다.	sa	sa	sa	100%
19. 안녕하세요. 반갑습니다.	sa	sa	sa	100%
20. 안녕하세요. 반갑습니다.	sa	sa	sa	100%
21. 안녕하세요. 반갑습니다.	sa	sa	sa	100%
22. 안녕하세요. 반갑습니다.	sa	sa	sa	100%
23. 안녕하세요. 반갑습니다.	sa	sa	sa	100%
24. 안녕하세요. 반갑습니다.	sa	sa	sa	100%
25. 안녕하세요. 반갑습니다.	sa	sa	sa	100%
26. 안녕하세요. 반갑습니다.	sa	sa	sa	100%
27. 안녕하세요. 반갑습니다.	sa	sa	sa	100%
28. 안녕하세요. 반갑습니다.	sa	sa	sa	100%
29. 안녕하세요. 반갑습니다.	sa	sa	sa	100%
30. 안녕하세요. 반갑습니다.	sa	sa	sa	100%
31. 안녕하세요. 반갑습니다.	sa	sa	sa	100%
32. 안녕하세요. 반갑습니다.	sa	sa	sa	100%
33. 안녕하세요. 반갑습니다.	sa	sa	sa	100%
34. 안녕하세요. 반갑습니다.	sa	sa	sa	100%
35. 안녕하세요. 반갑습니다.	sa	sa	sa	100%
36. 안녕하세요. 반갑습니다.	sa	sa	sa	100%
37. 안녕하세요. 반갑습니다.	sa	sa	sa	100%
38. 안녕하세요. 반갑습니다.	sa	sa	sa	100%
39. 안녕하세요. 반갑습니다.	sa	sa	sa	100%
40. 안녕하세요. 반갑습니다.	sa	sa	sa	100%
41. 안녕하세요. 반갑습니다.	sa	sa	sa	100%
42. 안녕하세요. 반갑습니다.	sa	sa	sa	100%
43. 안녕하세요. 반갑습니다.	sa	sa	sa	100%
44. 안녕하세요. 반갑습니다.	sa	sa	sa	100%
45. 안녕하세요. 반갑습니다.	sa	sa	sa	100%
46. 안녕하세요. 반갑습니다.	sa	sa	sa	100%
47. 안녕하세요. 반갑습니다.	sa	sa	sa	100%
48. 안녕하세요. 반갑습니다.	sa	sa	sa	100%
49. 안녕하세요. 반갑습니다.	sa	sa	sa	100%
50. 안녕하세요. 반갑습니다.	sa	sa	sa	100%

XML data format

```

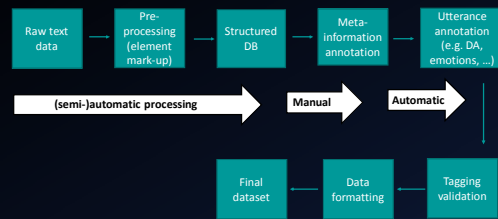
<?xml version="1.0" encoding="UTF-8" standalone="yes" ?>
<document type="Korpus" >
  <turn type="utterance" >
    <id type="utterance" value="1" />
    <text type="utterance" value="안녕하세요. 반갑습니다." />
    <category type="utterance" value="sa" />
  </turn>
  <turn type="utterance" >
    <id type="utterance" value="2" />
    <text type="utterance" value="안녕하세요. 반갑습니다." />
    <category type="utterance" value="sa" />
  </turn>
  <turn type="utterance" >
    <id type="utterance" value="3" />
    <text type="utterance" value="안녕하세요. 반갑습니다." />
    <category type="utterance" value="sa" />
  </turn>
  <turn type="utterance" >
    <id type="utterance" value="4" />
    <text type="utterance" value="안녕하세요. 반갑습니다." />
    <category type="utterance" value="sa" />
  </turn>
  <turn type="utterance" >
    <id type="utterance" value="5" />
    <text type="utterance" value="안녕하세요. 반갑습니다." />
    <category type="utterance" value="sa" />
  </turn>
  <turn type="utterance" >
    <id type="utterance" value="6" />
    <text type="utterance" value="안녕하세요. 반갑습니다." />
    <category type="utterance" value="sa" />
  </turn>
  <turn type="utterance" >
    <id type="utterance" value="7" />
    <text type="utterance" value="안녕하세요. 반갑습니다." />
    <category type="utterance" value="sa" />
  </turn>
  <turn type="utterance" >
    <id type="utterance" value="8" />
    <text type="utterance" value="안녕하세요. 반갑습니다." />
    <category type="utterance" value="sa" />
  </turn>
  <turn type="utterance" >
    <id type="utterance" value="9" />
    <text type="utterance" value="안녕하세요. 반갑습니다." />
    <category type="utterance" value="sa" />
  </turn>
  <turn type="utterance" >
    <id type="utterance" value="10" />
    <text type="utterance" value="안녕하세요. 반갑습니다." />
    <category type="utterance" value="sa" />
  </turn>
  <turn type="utterance" >
    <id type="utterance" value="11" />
    <text type="utterance" value="안녕하세요. 반갑습니다." />
    <category type="utterance" value="sa" />
  </turn>
  <turn type="utterance" >
    <id type="utterance" value="12" />
    <text type="utterance" value="안녕하세요. 반갑습니다." />
    <category type="utterance" value="sa" />
  </turn>
  <turn type="utterance" >
    <id type="utterance" value="13" />
    <text type="utterance" value="안녕하세요. 반갑습니다." />
    <category type="utterance" value="sa" />
  </turn>
  <turn type="utterance" >
    <id type="utterance" value="14" />
    <text type="utterance" value="안녕하세요. 반갑습니다." />
    <category type="utterance" value="sa" />
  </turn>
  <turn type="utterance" >
    <id type="utterance" value="15" />
    <text type="utterance" value="안녕하세요. 반갑습니다." />
    <category type="utterance" value="sa" />
  </turn>
  <turn type="utterance" >
    <id type="utterance" value="16" />
    <text type="utterance" value="안녕하세요. 반갑습니다." />
    <category type="utterance" value="sa" />
  </turn>
  <turn type="utterance" >
    <id type="utterance" value="17" />
    <text type="utterance" value="안녕하세요. 반갑습니다." />
    <category type="utterance" value="sa" />
  </turn>
  <turn type="utterance" >
    <id type="utterance" value="18" />
    <text type="utterance" value="안녕하세요. 반갑습니다." />
    <category type="utterance" value="sa" />
  </turn>
  <turn type="utterance" >
    <id type="utterance" value="19" />
    <text type="utterance" value="안녕하세요. 반갑습니다." />
    <category type="utterance" value="sa" />
  </turn>
  <turn type="utterance" >
    <id type="utterance" value="20" />
    <text type="utterance" value="안녕하세요. 반갑습니다." />
    <category type="utterance" value="sa" />
  </turn>
  <turn type="utterance" >
    <id type="utterance" value="21" />
    <text type="utterance" value="안녕하세요. 반갑습니다." />
    <category type="utterance" value="sa" />
  </turn>
  <turn type="utterance" >
    <id type="utterance" value="22" />
    <text type="utterance" value="안녕하세요. 반갑습니다." />
    <category type="utterance" value="sa" />
  </turn>
  <turn type="utterance" >
    <id type="utterance" value="23" />
    <text type="utterance" value="안녕하세요. 반갑습니다." />
    <category type="utterance" value="sa" />
  </turn>
  <turn type="utterance" >
    <id type="utterance" value="24" />
    <text type="utterance" value="안녕하세요. 반갑습니다." />
    <category type="utterance" value="sa" />
  </turn>
  <turn type="utterance" >
    <id type="utterance" value="25" />
    <text type="utterance" value="안녕하세요. 반갑습니다." />
    <category type="utterance" value="sa" />
  </turn>
  <turn type="utterance" >
    <id type="utterance" value="26" />
    <text type="utterance" value="안녕하세요. 반갑습니다." />
    <category type="utterance" value="sa" />
  </turn>
  <turn type="utterance" >
    <id type="utterance" value="27" />
    <text type="utterance" value="안녕하세요. 반갑습니다." />
    <category type="utterance" value="sa" />
  </turn>
  <turn type="utterance" >
    <id type="utterance" value="28" />
    <text type="utterance" value="안녕하세요. 반갑습니다." />
    <category type="utterance" value="sa" />
  </turn>
  <turn type="utterance" >
    <id type="utterance" value="29" />
    <text type="utterance" value="안녕하세요. 반갑습니다." />
    <category type="utterance" value="sa" />
  </turn>
  <turn type="utterance" >
    <id type="utterance" value="30" />
    <text type="utterance" value="안녕하세요. 반갑습니다." />
    <category type="utterance" value="sa" />
  </turn>
  <turn type="utterance" >
    <id type="utterance" value="31" />
    <text type="utterance" value="안녕하세요. 반갑습니다." />
    <category type="utterance" value="sa" />
  </turn>
  <turn type="utterance" >
    <id type="utterance" value="32" />
    <text type="utterance" value="안녕하세요. 반갑습니다." />
    <category type="utterance" value="sa" />
  </turn>
  <turn type="utterance" >
    <id type="utterance" value="33" />
    <text type="utterance" value="안녕하세요. 반갑습니다." />
    <category type="utterance" value="sa" />
  </turn>
  <turn type="utterance" >
    <id type="utterance" value="34" />
    <text type="utterance" value="안녕하세요. 반갑습니다." />
    <category type="utterance" value="sa" />
  </turn>
  <turn type="utterance" >
    <id type="utterance" value="35" />
    <text type="utterance" value="안녕하세요. 반갑습니다." />
    <category type="utterance" value="sa" />
  </turn>
  <turn type="utterance" >
    <id type="utterance" value="36" />
    <text type="utterance" value="안녕하세요. 반갑습니다." />
    <category type="utterance" value="sa" />
  </turn>
  <turn type="utterance" >
    <id type="utterance" value="37" />
    <text type="utterance" value="안녕하세요. 반갑습니다." />
    <category type="utterance" value="sa" />
  </turn>
  <turn type="utterance" >
    <id type="utterance" value="38" />
    <text type="utterance" value="안녕하세요. 반갑습니다." />
    <category type="utterance" value="sa" />
  </turn>
  <turn type="utterance" >
    <id type="utterance" value="39" />
    <text type="utterance" value="안녕하세요. 반갑습니다." />
    <category type="utterance" value="sa" />
  </turn>
  <turn type="utterance" >
    <id type="utterance" value="40" />
    <text type="utterance" value="안녕하세요. 반갑습니다." />
    <category type="utterance" value="sa" />
  </turn>
  <turn type="utterance" >
    <id type="utterance" value="41" />
    <text type="utterance" value="안녕하세요. 반갑습니다." />
    <category type="utterance" value="sa" />
  </turn>
  <turn type="utterance" >
    <id type="utterance" value="42" />
    <text type="utterance" value="안녕하세요. 반갑습니다." />
    <category type="utterance" value="sa" />
  </turn>
  <turn type="utterance" >
    <id type="utterance" value="43" />
    <text type="utterance" value="안녕하세요. 반갑습니다." />
    <category type="utterance" value="sa" />
  </turn>
  <turn type="utterance" >
    <id type="utterance" value="44" />
    <text type="utterance" value="안녕하세요. 반갑습니다." />
    <category type="utterance" value="sa" />
  </turn>
  <turn type="utterance" >
    <id type="utterance" value="45" />
    <text type="utterance" value="안녕하세요. 반갑습니다." />
    <category type="utterance" value="sa" />
  </turn>
  <turn type="utterance" >
    <id type="utterance" value="46" />
    <text type="utterance" value="안녕하세요. 반갑습니다." />
    <category type="utterance" value="sa" />
  </turn>
  <turn type="utterance" >
    <id type="utterance" value="47" />
    <text type="utterance" value="안녕하세요. 반갑습니다." />
    <category type="utterance" value="sa" />
  </turn>
  <turn type="utterance" >
    <id type="utterance" value="48" />
    <text type="utterance" value="안녕하세요. 반갑습니다." />
    <category type="utterance" value="sa" />
  </turn>
  <turn type="utterance" >
    <id type="utterance" value="49" />
    <text type="utterance" value="안녕하세요. 반갑습니다." />
    <category type="utterance" value="sa" />
  </turn>
  <turn type="utterance" >
    <id type="utterance" value="50" />
    <text type="utterance" value="안녕하세요. 반갑습니다." />
    <category type="utterance" value="sa" />
  </turn>

```

Scalability and updatability

- Automatic corpus processing (from raw to structured DB)
- Semi-automatic annotation (meta-information and utterance annotation)
- Annotated data could be used a training data for new updated data.

Corpus data construction



Applications

- Dialogue systems training
- Dialogue analysis research (discourse analysis)
- Spoken language research
- Language education – dialogue strategy

Limits

- No parallel audio or video data
- Sometimes unnatural conversation situations (mainly affected by the genres of drama and movie)

Acknowledgments

- This work was supported by the National Research Foundation of Korea Grant funded by the Korean Government(NRF-2017M3C4A7068186)

References

- Forchini, P. (2009, July). Spontaneously reloaded: American face-to-face and movie conversation compared. In *Corpus Linguistics 2009: Abstracts The 30th Corpus Linguistics Conference 20-23 July 2009, Liverpool* (p. 118).
- Quaglio, P. (2009). *Television dialogue: The sitcom Friends vs. natural conversation* (Vol. 36). John Benjamins Publishing.
- Serban, I. V., Lowe, R., Henderson, P., Charlin, L., & Pineau, J. (2015). A survey of available corpora for building data-driven dialogue systems. *arXiv preprint arXiv:1512.05742*.
- Serban, I. V., Sordani, A., Bengio, Y., Courville, A. C., & Pineau, J. (2016, February). Building End-To-End Dialogue Systems Using Generative Hierarchical Neural Network Models. In *AAAI* (Vol. 16, pp. 3776-3784).
- Sordani, A., Galley, M., Auli, M., Brockett, C., Ji, Y., Mitchell, M., ... & Dolan, B. (2015). A neural network approach to context-sensitive generation of conversational responses. *arXiv preprint arXiv:1506.06714*.
- Traum, D. R. (2000). 20 questions on dialogue act taxonomies. *Journal of semantics*, 17(1), 7-30.
- Webb, S. (2010). A corpus driven study of the potential for vocabulary learning through watching movies. *International Journal of Corpus Linguistics*, 15(4), 497-519.

References

- Forchini, P. (2009, July). Spontaneously reloaded: American face-to-face and movie conversation compared. In *Corpus Linguistics 2009: Abstracts The 30th Corpus Linguistics Conference 20-23 July 2009, Liverpool* (p. 118).
- Quaglio, P. (2009). *Television dialogue: The sitcom Friends vs. natural conversation* (Vol. 36). John Benjamins Publishing.
- Serban, I. V., Lowe, R., Henderson, P., Charlin, L., & Pineau, J. (2015). A survey of available corpora for building data-driven dialogue systems. *arXiv preprint arXiv:1512.05742*.
- Serban, I. V., Sordani, A., Bengio, Y., Courville, A. C., & Pineau, J. (2016, February). Building End-To-End Dialogue Systems Using Generative Hierarchical Neural Network Models. In *AAAI* (Vol. 16, pp. 3776-3784).
- Sordani, A., Galley, M., Auli, M., Brockett, C., Ji, Y., Mitchell, M., ... & Dolan, B. (2015). A neural network approach to context-sensitive generation of conversational responses. *arXiv preprint arXiv:1506.06714*.
- Traum, D. R. (2000). 20 questions on dialogue act taxonomies. *Journal of semantics*, 17(1), 7-30.
- Webb, S. (2010). A corpus driven study of the potential for vocabulary learning through watching movies. *International Journal of Corpus Linguistics*, 15(4), 497-519.

한국어의 ‘-ㅁ’과 ‘-기’에 대한 코퍼스 기반 연구

조기현/이용훈 (군산대학교/충남대학교)

1. 서론

A국어의 명사화소 -ㅁ/-기에 대한 연구는 국어에 대한 연구의 초반기부터 계속 연구되어 오고 있는 주제들 중의 하나로 그동안 이에 대한 많은 연구들이 있었다 (최현배, 1937; 장석진, 1966; 임홍빈, 1974; 채완, 1979; 홍종선, 1983; 김명광, 2004; 이상욱, 2004). 이러한 선행연구들은 대부분 -ㅁ으로 만들어진 명사들이 -기로 만들어진 명사들과 분포와 통사, 그리고 의미적인 측면에서 어떻게 다른가 하는 문제들에 대한 이론적 분석들이 많았다.

본 논문에서는 이러한 선행연구들과는 달리 코퍼스를 기반으로 -ㅁ/-기를 분석하려고 한다. 우선 우리는 세종구문분석 코퍼스에서 -ㅁ/-기가 들어간 문장들을 모두 추출하였다. 세종 원시코퍼스는 형태소분석이 되어있지 않아 -ㅁ/-기를 추출하기 어려웠고, 분석과정에서 통사정보를 활용하면 추출이 용이하였기 때문에, 본 연구에서는 세종구문분석 코퍼스에서 -ㅁ/-기가 들어간 문장들을 모두 추출하였다. 그다음, 영어에서의 선행연구들을 참조하여 20여개의 언어학적 요소들에 대하여 주석을 붙였다. 마지막으로 이렇게 모아진 문장들을 회귀분석을 사용하여 분석하였다. 이러한 분석을 우리는 어떠한 언어학적 요소들이 -ㅁ/-기의 선택에 관여하며 또한 코퍼스에서 추출된 데이터들이 얼마나 선행연구들의 주장들을 얼마나 뒷받침하고 있는가를 실증적 자료를 통하여 분석해 보고자 한다.

2. 선행연구

지금까지 ‘-ㅁ’과 ‘-기’에 대한 연구는 꾸준히 진행되어 왔으며, 김영광(2004)는 다음과 같이 정리하고 있다.

	-ㅁ	-기
최현배 (1937)	관념적	구체적 진행
장석진 (1966)	추상적·질적	구체적·양적
김재운 (1973)	내면적·합축적	외면적·구상적
임홍빈 (1974)	[+존재], [+대상화]	[-존재], [-대상화]
양동휘 (1975)	사실성	기대성
채완 (1979)	구체적 1회 사건	계속적 반복적 일반화된 개념
심재기 (1980)	[+결정성], [+실체성]	[-결정성], [-실체성]
홍종선 (1983)	순간성, 현장성, 과거성	일반화된 개념, 미래성
김남길 (1984)	사실성	비사실성
우형식 (1987)	개별적, 구체적, 실제적	일반적, 추상적, 가상적

<표 1> ‘-ㅁ’과 ‘-기’에 대한 선행 연구

이밖에도 ‘-ㅁ’과 ‘-기’의 분포는 용언의 품사, 복합용언인지 아닌지의 여부, 한자인지 아닌지의 여부, ‘-하’나 ‘-되’인지 아닌지의 여부, verbal noun인지 아닌지의 여부, 능동형인지 피동형인지의 여부 등에 영향을 받을 수 있다.

3. 연구 방법

3.1. 코퍼스와 분석에 사용된 팩터들

본 연구에서는 세종코퍼스에 포함되어 있는 여러 가지 코퍼스들 중에서 세종통사분석코퍼스(433,839 어절; 43,828 문장)를 선택하였다. 이 코퍼스에서 먼저 ‘-ㅁ’과 ‘-기’가 들어간 문장들을 형태소 정보를 활용하여 모두 추출하였다. 그러나 추출된 문장의 양이 사람이 주석을 첨가하기에는 너무 많은 양이어서, 추출된 문장들 중에서 임의로 1,000문장을 추출하였다. 이 1,000문장 중에는 lexical nominalization을 포함하는 문장들은 제외하고 structural nominalization을 포함하는 문장들만을 포함시켰다. 그리고 이렇게 추출된 각각의 문장들에 대하여 다음과 같은 언어학적 팩터들을 하나씩 하나씩 코딩하였다.

Type	Variable (ID Tag)	Comment	Levels (ID Tag Levels)
Complement Verb	CVVerb	용언의 Yale Romanization	
	CVPOS	용언의 품사	a, v
	CVComplex	복합용언인지 아닌지의 여부	yes, no
	CVSinoKorean	한자인지 아닌지의 여부	yes, no
	CVHaToy	‘-하’나 ‘-되’인지 아닌지의 여부	ha, toy, no
	CVVerbal	verbal noun인지 아닌지의 여부	yes, no
	CVVoice	능동형인지 피동형인지의 여부	active, passive
	CVTransitivity	용언의 valency	intransitive, transitive, ditransitive
	CVCompForm	용언의 보어 혹은 목적어의 형태	na, np, ap, cp, ...
	CVType	용언의 Vendler Type	state, process, accomp., achievement
CVSem	용언의 의미타입	abstract, concrete	
Matrix Verb	MVVerb	용언의 Yale Romanization	
	MVPOS	용언의 품사	a, v
	MVComplex	복합용언인지 아닌지의 여부	yes, no
	MVSinoKorean	한자인지 아닌지의 여부	yes, no
	MVHaToy	‘-하’나 ‘-되’인지 아닌지의 여부	ha, toy, no
	MVVerbal	verbal noun인지 아닌지의 여부	yes, no
	MVVoice	능동형인지 피동형인지의 여부	active, passive
	MVTransitivity	용언의 valency	intransitive, transitive, ditransitive
	MVCompForm	용언의 보어 혹은 목적어의 형태	na, np, ap, cp, ...
	MVType	용언의 Vendler Type	state, process, accomp., achievement
MVSem	용언의 의미타입	abstract, concrete	

<표 2> 분석에 사용된 언어학적 팩터들

코퍼스에서 추출된 문장들에 대하여 <표 2>의 모든 팩터들에 대한 주석을 달았다.

3.2. 통계분석

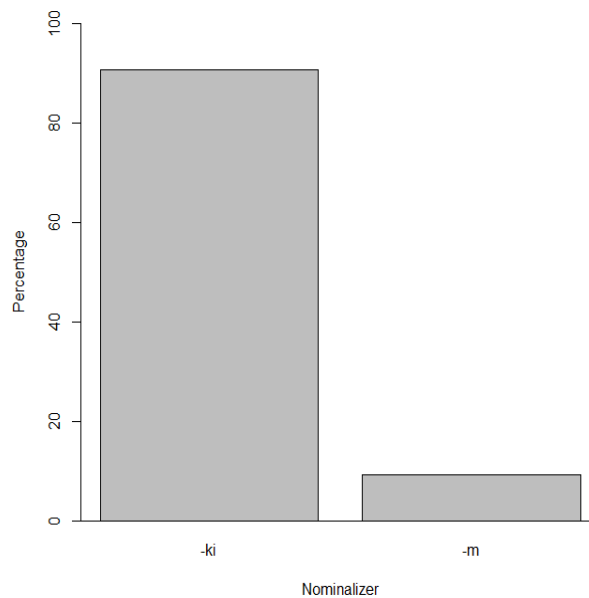
코퍼스분석에 사용된 언어학적 팩터들이 약 20개정도나 되기 때문에, 이 코퍼스자료들을 분석하기 위해서 분산분석(ANOVA)을 사용하는 것은 비효율적이다. 따라서, 본 연구에서는 분산분석이 아닌 회귀분

석, 그 중에서도 로지스틱 회귀분석(logistic regression)을 사용하였다. 또한 모든 통계분석에는 R이 사용되었다.

4. 분석결과

4.1. 기본분석 결과

‘-ㅁ’과 ‘-기’의 기본적인 통계분석결과는 다음과 같다.



<그림 1> ‘-ㅁ’과 ‘-기’의 분포

이 그래프가 보여주듯이 전체 문장들 중에서 10% 정도의 문장들만이 ‘-ㅁ’을 사용하였으며 나머지 90% 정도의 문장들이 ‘-기’를 사용하였다.

4.2. 회귀분석 결과

세종통사분석코퍼스에서 추출한 문장들에서 ‘-ㅁ’과 ‘-기’의 분포에 대한 로지스틱 회귀분석의 결과는 다음과 같다.

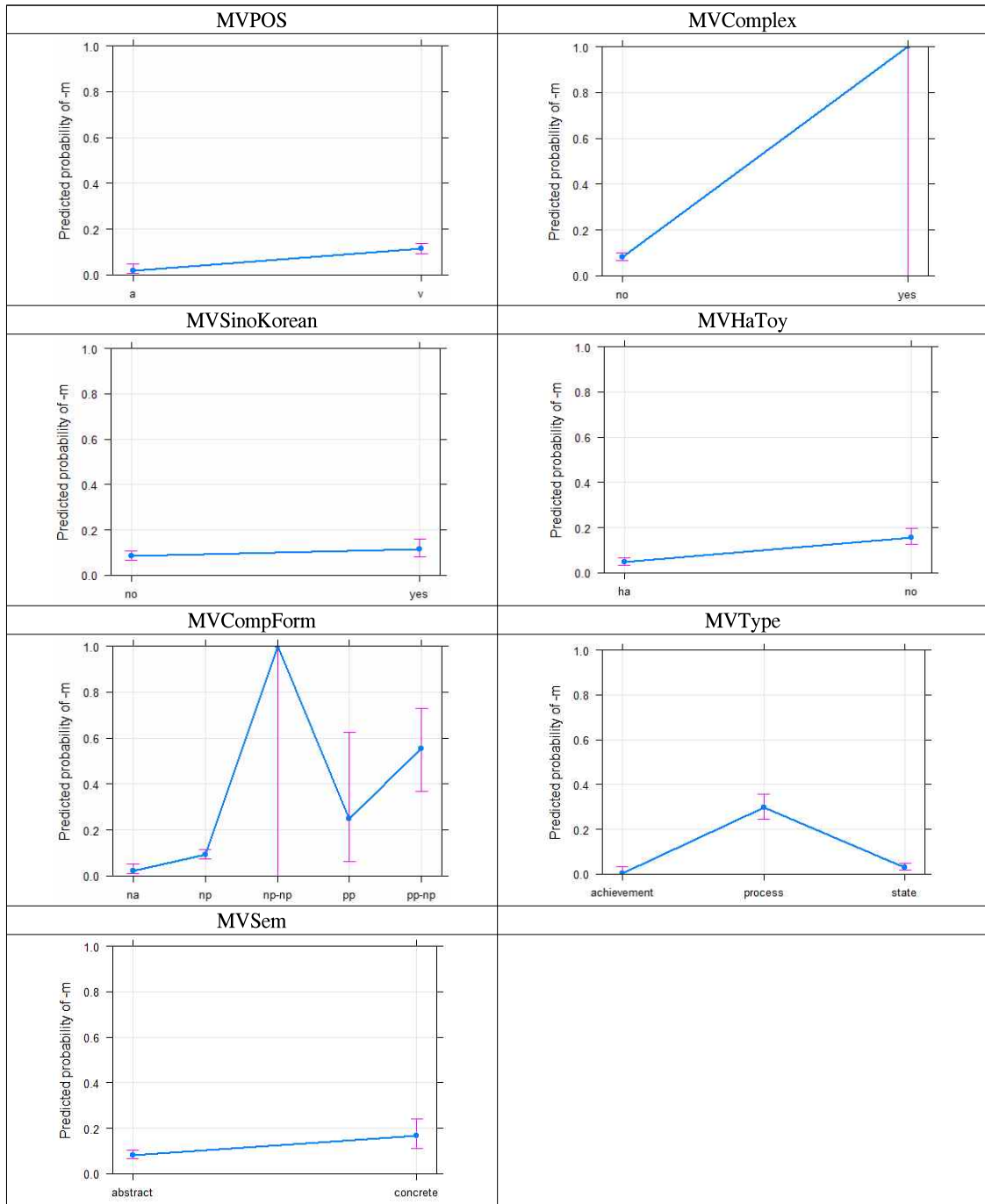
	<i>df</i>	Deviance	AIC	LRT	<i>p</i>
<none>	198.91	268.91			
CVPOS	1	200.82	268.82	1.905	.168
CVComplex	1	198.94	266.94	0.028	.867
CVSinoKorean	1	199.21	267.21	0.296	.586
CVHaToy	2	200.30	266.30	1.391	.499
CVVerbal	1	199.12	267.12	0.214	.644
CVVoice	1	199.23	267.23	0.321	.571
CVTransitivity	2	203.35	269.35	4.439	.109
CVCompForm	4	205.16	267.16	6.251	.181
CVType	3	203.56	267.56	4.654	.199
CVSem	3	203.71	267.71	4.797	.187
MVPOS	1	206.65	274.65	7.742	.005 **
MVComplex	1	218.33	286.33	19.417	<.001 ***
MVSinoKorean	1	210.03	278.03	11.121	<.001 **
MVHaToy	1	205.39	273.39	6.477	.011 *
MVVerbal	1	199.74	267.74	0.830	.362
MVVoice	1	198.91	266.91	0.000	.999
MVTransitivity	1	198.92	266.92	0.006	.940
MVCompForm	3	209.45	273.45	10.538	.015 *
MVType	2	247.48	313.48	48.568	<.001 ***
MVSem	2	214.45	280.45	15.542	<.001 ***

<표 3> 로지스틱 회귀분석의 결과

이 표를 보면 알 수 있듯이, 전체 20개의 팩터들 중에 7개의 팩터들만의 ‘-ㅁ’과 ‘-기’의 분포에 유의미하게 관여하고 있음을 알 수 있다.

4.3. 그래프를 활용한 분석

<표 2>에서 유의미하다고 판단된 7개의 팩터들을 effect plot으로 분석해보면 다음과 같다.



<그림 2> Effect Plot의 분석 결과

이 그림들에서 알 수 있듯이, 한국어에서 ‘-ㅁ’과 ‘-기’의 선택은 이들을 보거나 목적으로 취하는 주절 동사의 성격에 따라 결정됨을 알 수 있다.

5. 결론

본 논문에서는 세종통사분석코퍼스에서 먼저 ‘-ㅁ’과 ‘-기’가 들어간 1,000 문장들을 형태소 정보를 활용하여 모두 추출하였다. 그리고 이렇게 추출된 각각의 문장들에 대하여 20여개의 언어학적 팩터들을 하나씩 하나씩 코딩하였다. 그런 다음, 로지스틱 회귀분석을 사용하여 통계적으로 분석하였다.

통계분석의 결과, 전체 문장들 중에서 10% 정도의 문장들만이 ‘-ㅁ’을 사용하고 있으며 나머지 90% 정도의 문장들이 ‘-기’를 사용하고 있음을 알 수 있었으며, 20여개의 언어학적 팩터들 중에서 7개의 팩터들만의 ‘-ㅁ’과 ‘-기’의 분포에 유의미하게 관여하고 있음을 알 수 있었다.

참고문헌

- 김명광 (2004), 「국어접사 -음, -기에 의한 단어형성 연구」. 서강대학교 박사논문.
이상욱 (2004), 「-음, -기 명사형성 단어화에 대한 연구」. 서울대학교 석사논문.
임홍빈 (1974), “명사화의 의미특성에 대하여”, 「국어학」 2, 83-104.
장석진 (1966), Some Remarks on Korean Nominalization. 「어학연구」 2(1), 18-31.
채완 (1979), “명사화소 ’-기’에 대하여”, 「국어학」 8, 95-107.
최현배 (1937), 「우리말본」. 서울: 정음사.
홍중선 (1983), “명사화 어미의 변천”, 「국어국문학」 89, 31-51.

‘기는’ 구문의 정보 분할과 분열문

김건희(서원대)

1. 서론

(1)에 대한 대답은 (2), (3)으로 다양하게 나타날 수 있다.

- (1) 영이가 철수에게 뭘 줬어?
- (2) ㄱ. 신발 (영이가 철수에게 신발 줬어.)
 나. 영이가 철수에게 준 것은 신발이야.
- (3) ㄱ. 영이가 철수에게 주기는 신발을 줬어.
 나. 영이가 주기는 철수에게 주기는 신발을 줬어.

(2ㄱ)은 질문에 주어진 요소를 생략시키고 질문에 대한 대답인 초점, ‘신발’만 제시하여 돋보인다.¹⁾ (2나)은 초점의 대표적인 통사적 기제인 분열문으로 제시되어 ‘신발’이 돋보인다. (3)의 경우, 어순변화, 형태적 표지 결합(‘-기’, ‘-는’)이 나타나 (2)에 비해 더 돋보인다. 이처럼 정보 구조 논의에서 초점의 전형적인 화용론적 기능은 질문의 의문사(wh-) 요소에 대응하는 대답 요소인 ‘초점’ 자신을 돋보이게 (highlight) 하는 것이다.(Paul 1880) 이러한 Paul(1880)의 돋보임/두드러짐(highlighting)은 잘 알려진 Chafe(1976)의 정보 포장과 연계된다. 곧 (2), (3) 모두 질문에 대한 대답인 초점 ‘신발’을 나타내면서 정보 자체는 동일하지만 다양한 정보 포장(information packaging, how information is presented, in contrast to the information itself, Chafe(1976:28))으로 초점이 돋보이게 나타난다는 것을 알 수 있다. 정보 그 자체는 변함이 없지만 다양한 형태, 통사적 기제로 정보 포장에 있어서 차이를 보이는 것이다.

본고에서는 ‘(2) 생략, 일반 분열문’의 정보 포장보다 담화적 돌출성(discourse prominence)이 나타나 (3) ‘기는’ 구문에 대해서 살펴보겠다. ‘기는’ 구문에 대해서는 일찍이 동사(용언) 반복 구문(No(1998), 최기용(2002)), 술부 분열문(Lee(1995), 이정훈(2013))으로 제시하며 생성 문법적 관점에서 다루었고 ‘술어의 주제화, 초점화’(이수연(2014), 이금희(2015))로 정보 구조적 관점에서 다루거나 문법화 관점에서 다루기도 하였다.(구현정(2012)) 생성 문법적 관점이나 문법화 관점 모두 정보 구조적 측면을 어느 정도 전제로 하고 있어 ‘기는’ 구문은 주격 중출 구문과 함께 정보 구조 논의가 주목받기 훨씬 이전부터 일찍이 정보 구조적 측면에서 논의된 대표적인 구문이다. ‘기는’ 구문에 대해 구체적으로 다음과 같은 논제를 중심으로 살펴보겠다.

첫째 기존 연구에서는 ‘기는’ 구문을 ‘술부 분열문’으로 제시하거나 ‘화제화/주제화 구문’으로 제시하였는데 특히 ‘(...)V/A+기는’을 ‘화제’ 혹은 ‘주제’로 보는 논의도 적지 않다.²⁾ 그러나 ‘명사(류)+-은/는’

1) 생략(ellipsis)은 주어집성을 표시하는 극단적인 형태로도 제시되지만(Krifka 2008:263), 범언어적으로 생략은 주어진 요소이나 배경적 요소(given or background)인 V(P)를 표시하는 보편적인 정보 포장 전략이다.Zimmermann(2016:334)

2) 본고에서는 ‘-기’에 ‘-는’이 결합한 ‘기는’ 구문에 대해 다루지만 일찍이 임홍빈(1974:92)에서는 명사화소 ‘-기’에

이 문두에 제시되는 일반적인 문장과 달리 용언이 명사화된 형태에 ‘-는’이 결합되어 문두에 제시되고 ‘초점’이 따로 존재한다는 점에서 ‘기는’ 구문의 정보 분할에 대해서는 좀 더 정교하게 논의되어야 한다.

둘째 본고에서는 Drubig(2000)에 기대어 ‘기는’ 구문을 계사가 없고 술어를 중심으로 한 단위(V, VP, S)가 명사화되어 전제로 나타나는, 소위 제한된 분열문(reduced cleft)(Drubig(2000:71)으로 제시한다. 이를 위해 Dik(1981), Dik(1997)에서 제시한 초점 유형인 보충 초점(completive focus)을 정교하게 발전시킨 Drubig(2000)의 분열문 논의를 자세히 살펴보고자 한다.

셋째 정보 구조 논의에서는 어떤 질문이 상정되느냐에 따라 정보 분할이 달라지는 등 질문의 상정이 중요한 부분을 차지하는데(임동훈(2012), 전영철(2013), 박철우(2015), 함병호(2016), 김민국(2016), Velleman and beaver(2016:92) 등 참조) 문제는 이러한 질문이 명시적으로 제시되지 않아 자의적으로 상정될 수도 있다는 점이다. 본고에서는 ‘기는’ 구문이 전제로 환원되는 ‘질문’과 초점으로 나타나는 ‘대답’ 쌍으로 재분석될 수 있음을 제시한다. ‘기는’ 구문은 일찍이 Carlson(1983)에서 제시한, 소위 스스로 답하는 질문(self-answered question)에 해당되어 질문이 전제의 형태로 명시적으로 드러난다는 점에서 분석 고찰의 의의가 있다.

이러한 논제를 토대로 본고에서 다룰 내용은 다음과 같다. 2장에서는 문두에 나타나는 ‘(...)V/A+기는’의 정보 단위에 대해 살펴 보면서 ‘기는’ 구문의 특이한 정보 분할에 대해 논의하겠다. 3장에서는 앞서 ‘(...)V/A+기는’의 정보 단위를 초점에 대응되는 ‘전제’로 본 2장을 토대로 전제의 정보적 가치에 대해 살펴 보겠다. 4장에서는 ‘기는’ 구문을 분열문으로 보는 것에 대해 Dik(1981), Dik(1997)의 보충 초점과 Drubig(2000)의 분열문 논의를 토대로 살펴 보겠다. 5장 결론에서는 논의 결과를 요약하고 남은 문제를 제시하겠다.

참고문헌

- 구현정(2012), “‘-기는’ 구문의 문법화”, 「담화와 인지」 19-3, 담화·인지학회, 1-28.
- 김건희(2014), “선행절에 어미 ‘-지’가 결합된 절이 나타나는 문장에 대한 연구”, 「한글」 303, 한글학회, 115-146.
- 김건희(2018), “주격중출 구문에 대한 일고찰 -정보구조와 문장 확대 구성 관점을 중심으로-”, 「우리말글」 76, 우리말글학회, 1-33.
- 김민국(2016), “한국어의 격조사와 초점-주격과 대격을 중심으로-”, 「국어학」 78, 국어학회, 165-219.
- 박철우(2015), “보조사의 기능과 정보구조”, 「국어학」 73, 국어학회, 269-307.
- 유은정(2011), “스페인어 술부 분열 구문과 한국어 ‘[V1-기는... V2-...]’ 구문에 대한 비교 연구와 스페인어 교육에서의 활용”, 「스페인어문학」 60, 한국스페인어학회, 103-127.
- 이금희(2015), “국어 서술어(구)의 초점화와 주제화에 관한 연구-‘기를, -기는’ 구성 중심으로”, 「한국어 의미학」 49, 한국어의미학회, 87-108.

‘-는, -야, -도’ 등이 결합하여 용언의 주제화를 이룬다는 것을 제시하였고(구현정(2012)에서 재인용) 이후 임흥빈(2007:533)에서는 ‘기를’이 문두에 제시되는 아래 예를 주제화 구문으로 제시하였다. 이러한 예시는 용언이 명사화되어 문두에 제시되는 ‘기는’ 구문과 같은 형상이다.(순이가 예쁘기는 아주 예쁘다)

가. 순이가 예쁘기를 아주 예쁘다.

나. 철수가 그 집에 가기를 너무 자주 그 집에 간다.

다. 순이가 얼굴을 보기를 하루에 열 번씩 얼굴을 본다.

라. 순이가 나를 알기를 아주 우습게 안다.

(임흥빈2007:533)

구현정(2012:4)에서는 ‘기는’ 구문이 주어진 정보, 전제된 정보를 주제로 제시한, 주제화한 구문이라고 하였고 이정훈(2014:218)에서도 ‘영이가 먹기는 떡을 먹었다.’와 같은 예시를 동사구 주제화 구문이라고 제시하였다.

- 이수연(2014), “동일 어간 반복 구문의 의미 특성 - ‘V-기는 V, V-기만 V’ 구문을 중심으로”, 「한국어 의미학」 46, 한국어의미학회, 141-161.
- 이정훈(2013), “‘V-기’ 반복 구문의 유형과 그 형성 동기 및 과정”, 「어문학」 122, 한국어문학회, 155-180.
- 이정훈(2014), “동사구 주제화 구문의 통사론”, 「국어학」 72, 국어학회, 205-233.
- 임동훈(2012), “‘은/는’과 중형의 의미 관계”, 「국어학」 64, 국어학회, 217-271.
- 임홍빈(1974), “명사화의 의미특성에 대하여”, 「국어학」 2, 국어학회, 83-104.
- 임홍빈(2007), 「한국어의 주제와 통사 분석: 주제 개념의 새로운 전개」, 서울대학교 출판부.
- 전영철(2006), “대조 화제와 대조 초점의 표지 ‘는’”, 「한글」 274, 한글학회, 171-200.
- 전영철(2013), “한국어의 제언문/정언문 구별과 정보구조”, 「국어학」 68, 국어학회, 99-133.
- 최기용(2002), “한국어 용언 반복 구문: 용언의 가시적 이동을 위한 또 하나의 근거”, 「생성문법연구」 12, 생성문법학회, 139-167.
- 최윤지(2016), “구어에서 유의미하게 나타나는 어떤 통사 구성에 대하여”, 「국어학」 79, 국어학회, 187-237.
- 함병호(2016), “정보구조의 분절 방법”, 「동악어문학」 67, 동악어문학회, 311-338.
- Büring, D.(2003), On d-trees, beans, and b-accent, *Linguistics and Philosophy* 26(5), 511 - 545
- Büring, D.(2016), (Contrastive) topic, in C. Féry, and S. Ishihara (eds.), *The Oxford Handbook of Information Structure*, Oxford University Press. 64-85.
- Carlson, L.(1983). *Dialogue Games: An Approach to Discourse Anaphora*, Dordrecht: Reidel.
- Chafe, W.(1976), Givenness, Contrastiveness, Definiteness, Subjects, in Charles Li. (ed.), *Topics and Point of View. Subject and Topic*, New York: Academic Press, 25-56.
- Chomsky, N.(1972) Deep structure, surface structure, and semantic interpretation, in Noam Chomsky (ed.), *Studies on Semantics in Generative Grammar*, The Hague: Mouton, 69 - 119.
- Dik, Simon C.(1997), *The Theory of Functional Grammar. Part I: The Structure of the Clause*, Berlin and New York, NY: Mouton de Gruyter.
- Dik, Simon C. et al.(1981), On the typology of focus phenomena. in Teun Hoekstra (ed.), *Perspectives on Functional Grammar*, Dordrecht: Foris, 41-74.
- Drubig, H. Bernhard(2000). Towards a typology of focus constructions. University of Tübingen, 1-113.
- Drubig, H. Bernhard(2003), Toward a Typology of Focus and Focus Constructions, *Linguistics* 41, 1-50.
- Erteschik-Shir, Nomi(1997). *The Dynamics of Focus Structure*, Cambridge University Press.
- Féry, C. and S. Ishihara(2016), Introduction, in C. Féry, and S. Ishihara (eds.), *The Oxford Handbook of Information Structure*, Oxford University Press. 1-18.
- Gundel, Jeanette. K.(1988), Universals of topic-comment structure. in M. Hammond, E. Moravczik and J. Wirth (eds.), *Studies in syntactic typology*, Amsterdam: John Benjamins, 209-239.
- Halliday, M. A. K.(1967), *Intonation and Grammar in British English*, The Hague; Paris: Mouton.
- Hengeveld, K. & Mackenzie, J. L.(2008), *Functional Discourse Grammar: A typologically-based theory of language structure*, Oxford University Press.
- Jackendoff, R.(1972), *Semantic Interpretation in Generative Grammar*, Cambridge, MA: MIT Press.
- Krifka, M.(2008), Basic Notions of Information Structure, *Acta Linguistica Hungarica* 55, 243-276.
- Lee, Chungmin.(1999), Topic, contrastive topic and focus: What's on our mind. Available at <<http://www.mind.sccs.chukyou.ac.jp/jcss/ICCS/99/olp/pti/pti.htm>>.
- Lee, J.-S.(1995), A Study on Predicate Clefting, *Studies in Generative Grammar* 5-2, 531-584.
- Manfred Krifka & Renate Musan(2012), Information structure: Overview and linguistic issues, in Manfred Krifka

- & Renate Musan (eds.), *The expression of information structure*, 1 - 44
- McCloskey, J.(1979), *Transformational Syntax and Model Theoretic Semantics*, Dordrecht: Reidel.
- Neeleman, Ad and Hans van de Koot(2008), Dutch Scrambling and the Nature of Discourse Templates, *Journal of Comparative Germanic Linguistics* 11, 137-189.
- Neeleman, Ad and Hans van de Koot(2016), Word order and information structure, in C. Féry, and S. Ishihara (eds.), *The Oxford Handbook of Information Structure*, Oxford University Press. 383-401.
- No Yongkyoon(1998), Negative morphemes in Korean: Evidence for a derivational treatment, in (ed.) E-J Baek, *Proceedings of the 6th International Conference of Korean Linguistics*, Toronto: Department of East Asian Studies, University of Toronto, 556-567.
- Paul, Hermann(1880/1995), *Prinzipien der Sprachgeschichte* (10th edn.). Tübingen: Niemeyer.
- Rochemont, M.(2016), Givenness, in C. Féry, and S. Ishihara (eds.), *The Oxford Handbook of Information Structure*, Oxford University Press, 41-63.
- Sæbø, Kjell Johan(2016), Information structure and Presupposition, in C. Féry, and S. Ishihara (eds.), *The Oxford Handbook of Information Structure*, Oxford University Press. 128-146.
- Sasse, Hans-Jürgen(1987), The Thetic/Categorical Distinction Revisited. *Linguistics* 25, 511-580.
- Velleman, L. and David Beaver(2016), Question-based models of information structure, in C. Féry, and S. Ishihara (eds.), *The Oxford Handbook of Information Structure*, Oxford University Press. 86-107.
- Wagner, M.(2012), Contrastive Topics Decomposed, *Semantics and Pragmatics* 5(8), 1-54.
- Zimmermann, M.(2016), Predicate focus, in C. Féry, and S. Ishihara (eds.), *The Oxford Handbook of Information Structure*, Oxford University Press, 314-335.

■ 개인 발표 3 ■

Epistemic modal, temporal orientation, and Korean bouletic verbs

Choi Mingyeong (SNU)

1. Introduction

This paper mainly deals with epistemic modal base and Korean bouletic verbs. I will look two parts connected with one another concerning the topic. First part is a review of previous studies about modal base and temporal orientation. Second part is modal analyses of Korean bouletic verbs using a traditional concept of modal base. Looking these parts at the same time can give us some intuition for issues. First, observation from bouletic verbs can answer the question whether the future orientation of epistemic modal base is possible, as recent paper Banerjee (2018) first introduced. Second, setting different modal bases might be an effective way to explain the different meaning between two bouletic verbs.

2. Previous studies of modal base and temporal orientation

2.1. Werner (2006)

Werner (2006) observed that temporal orientation of a modal is related to the modal base. He argues that epistemic modal has present and past orientation, while non-epistemic modal has a future orientation. His following sentences have an ambiguity between an attitude about future event and non-future event.

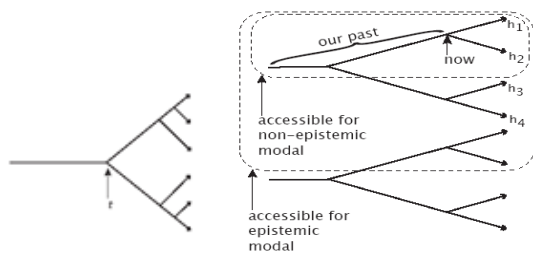
- (1) a. Jim might be late. (epistemic or stereotypical)
- b. That will be Jack. (epistemic or stereotypical)
- c. Jill may be seated. (epistemic or deontic)
- d. Janet must live in student housing. (epistemic or deontic)
- e. John can't be our representative. (epistemic or circumstantial)

(Werner 2006: (1-5))

He links a future orientation of a Non-epistemic modal to its 'forward-branching' modal base. That is, all worlds of the modal base have same history until speech time, and thereafter be distinguished (This is also called historical modal base). To satisfy a Disparity Principle, a modal claim can only be about a future event.

Non-future orientation of epistemic modal comes from the time branching model of the epistemic modal base, Disparity Principle, and Non-Disparity Principle. Non-Disparity Principle, which is added by Werner, requires that each

tree not to be distinguished from the speech time.



Circumstantial modal base(left) and epistemic modal base(right) (Portner 2009 : Figure 5.1-5.2)

2.2. Klecha (2015)

Klecha (2015)'s main observation is how an interpretation of non-past in finite clause embedded under an attitude verb is affected by a semantic meaning of the attitude verb. He first generalizes that epistemic modal has a non-future time interval worlds and non-epistemic modal has a future time interval worlds. In this section, I mainly focuses on the evidences he brought to claim about past and present orientation of epistemic modals.

First, when an epistemic modal in a sentence seems to have a future time preajacent, he suggests that it should be thought as a present schedule, not a future event. In these sentences as in (14), it is considered that covert schedule modal (Copley 2002) is embedded under a scope of an epistemic modal, *has to / must*. Then the epistemic modal is actually about a present schedule.

- (2) A: The ship leaves the Spanish harbor next Tuesday.
 B: It has to/ must reach Athens next Friday (then).
 (Klecha 2015: (16))
- (3) A: The doctor will be in London either next week or the following week.
 B: Actually, she will be at a conference in the US next week.
 A: Oh, she must be in London the following week then.
 (Klecha 2015: (17))

Next, he argues that a seemingly epistemic modal for future time preajacent is in fact a metaphysical modal, which is non-epistemic modal. For example, he has a different view for a *may* in (17c), which was treated as an epistemic modal by Condoravdi (2002).

He saw that *may* in (17c) cannot be used in this conext of (17), where the future event of John's meeting is already decided. He claims that this restriction results from the *may* as a circumstantial modal when it has a future proposition. Since the circumstance of this context is set that John meets a certain person, so circumstantial modal cannot satisfy a Disparity Principle, disallowing the final sentence of (17)

- (4) a. John is going to meet either the provost or the dean.
 b. It has been decided who he will meet, but I don't know who it is.
 c. He may meet the dean, and he may meet the provost.
 (Klecha 2015: (71))

2.3. Banerjee (2018)

Banerjee (2018) argues that epistemic modal base can have a future-orientation, showing specific situation where the modal base of future-shifted hope should be seen as an epistemic.

As for epistemic verb *think*, and epistemic modal *must*, which cannot be future-oriented, he finds this constraint of temporal orientation not on the epistemic modal base. For *must*, he makes the following principle, which can predict its non-future orientation when it is used as a epistemic modal.

- (5) Strong non-root necessity modals cannot be future-oriented (Banerjee 2018: (38))

As a contrast, he points out that weak necessity epistemic modal (having ordering source) should can be compatible with future orientation.

- (6) a. *The Haverhill train must arrive at Ballardvale at 19:35 (given it left North Station at 18:55).
 b. The Haverhill train should arrive at Ballardvale at 19:35 (given it left North Station at 18:55).

3. Previous studies of *hope* and *want*

3.1. *want* in Rubinstein (2017)

Heim (1992) first thought that *want* reflects the doxastic world of attitude holder, not the real world. However, Villalta (2008) shows that the modal base of *want* should not be doxastic. The following example is the evidence for such a claim.

- (7) a. I want to teach Tuesdays and Thursdays next semester.
 b. I believe that I will teach Tuesdays and Thursdays next semester if and only if I work hard now.
 c. Invalid inference: I want to work hard now.
 (Villalta 2008: 478)

The formulation that reflects Villalta's solution is given below. Here, comparing worlds is not on the doxastic alternatives, but on certain worlds related to the context.

- (8) $\llbracket \text{want} \rrbracket^{f,h}(q)(a)(w)$ is defined iff $h(a,w) = \text{DES}(a,w), \cap f(w) \cap \text{DOX}(a,w) \cap q \neq 0$,
and $\cap f(w) \cap \text{DOX}(a,w) \cap \neg q \neq 0$.

If defined, $\llbracket \text{want} \rrbracket^{f,h}(q)(a)(w) = 1$ iff $\cap f(w) \cap q <_{h(a,w)} \cap f(w) \cap \neg q$

(Rubinstein 2017: (14))

However even in this semantic of *want*, belief of the attitude holder concerns the interpretation, as *q*-world is required to be compatible with doxastic worlds to define want. According to Rubinstein, this assumption can also be questioned. This is because anyone can claim want *q* even if they believe that it is not possible.

3.2. *hope* in previous studies

Anand & hacquard (2013) labelled attitude verb *hope* as emotive doxastic. This is because on the one hand, it has a meaning of preference like *want*, and on the other hand, it has a representational meaning like *think* and *say*. emotive doxastic can contain epistemic modal in embedded clause due to the representational meaning, but only possibility modal due to the preference meaning.

Also, in discourses, this additional meaning of hope can be observed in contrast to want.

- (9) A: Kommt Peter heute?

comes Peter today

'Is Peter coming today?'

B: Ich hoffe/*will, dass er heute kommt.

I hope/*want that he today comes

'I hope/*want that he is coming today'

(Scheffler 2008)

- (10) It is raining.

a. # I hope it is raining/#that is what I hope.

b. "I want it to be raining/"that is what I want.

(Scheffler 2008)

- (11) It isn't raining.

a. # I hope it is raining/#that is not what I hope.

b. "I want it to be raining/"that is not what I want.

(Scheffler 2008)

Klecha(2015) says that *hope* has a doxastic modal base when its temporal orientation is non-future, and a circumstantial modal base when it has a future orientation. This is the claim that a type of modal base varies according

to the temporal orientation. Following example shows uncommon temporal orientation of *hope*, in contrast to other doxastic verbs that only has a non-future orientation. While non-past tense in embedded clause of *think* cannot refer to a future, that in embedded clause of *hope* can.

- (12) Martina hoped Carissa got pregnant.
a. Martina hopes Carissa gets pregnant.
b. Martina hopes Carissa got pregnant.
- (13) Martina thought Carissa got pregnant.
a. *Martina thinks Carissa gets pregnant.
b. Martina thinks Carissa got pregnant

However, I think his evidences are not that strong to claim such a temporal-modal correlation. In next chapter, I will suggest that *hope* has doxastic character even for the future event.

Banerjee(2018) says that *hope* has an epistemic modal even for a future-orientation. Through this new observation, he claims that the existing generalization of temporal-modal correlation should be corrected.

The virtual situation he sets is as follows. An evil queen opened a fencing competition, having in mind in secret willing to execute the winner (herself to be a best player). One of the participants, Yena, does not know the outcome of the victory. Then, the sentence *Yena hopes to win the competition*, which has an obvious future-orientation can be an appropriate report about Yena's state.

In this situation, if the modal base of the sentence were circumstantial, it would be a contradictory report. The relevant circumstance here is the situation under a queen's law, so every world Yena win is the world where Yena is executed.

Even though I agree with *hope* as an epistemic modal, I think it is questionable in Banerjee's explanation that the nature of circumstantial modal base really makes it contradictory report. To me, it is not certain whether queen's secret plan must be the relevant circumstance for the circumstantial modal claim. To avoid such a subtle problem, I would rather find its epistemic modal base in another aspect. That is, Yena's hope is not a natural one, but a secondary, belief-generated one. Yena's desire starts from a certain belief that winning the competition will bring her fame and wealth. In other words, the worlds relevant to her hope should basically be doxastic worlds.

4. Korean bouletic verbs *pala-/wenha-*

In the literature, while *want* was treated as having little relation with doxastic meaning (Villalta2008, Rubinstein 2017), *hope* was claimed to have it (Anand&Hacquard 2013, Banerjee 2018). But the examples cited above only have a present orientation (examples of scheffler 2008) and Banerjee (2018)'s example needs to be explained in more detail, I think. In this chapter, looking at Korean data, I will investigate more the phenomena of two kinds of bouletic verbs. These examples also contain complex constructions, not only different uses in situations. And in particular, I will

- (16) a. sikan-ul cay-n-ta-myen elmana toyl-ci
time-ACC measure-PRES-DECL-COND how become-NML
cimcakha-ki-to elyep-ta.
guess-NML-ADD hard-DECL
‘It is hard to even guess how long it will be if you measure the time’
- b. cikum ku-ka eti-ey iss-nun-ci
now he-NOM where-LOC exist-PRES-NML
chwuchukha-ki elyep-ta.
conjecture-NML hard-DECL
‘It is hard to guess where he is now’
- (17) a. pwukhan-un hayngpo-lul yeysangha-ki
North.Korea-TOP going-ACC anticipate-NML
elyew-un sangtay-ta.
hard-ADN opponent-DECL
‘North Korea is an opponent hard to anticipate its move’

It has a meaning that attitude holder has a difficulty thinking what is more likely in the situation. In many cases, these sentences have a interrogative like *who*, *which* or *what*. I think these properties come from insufficiency of ordering source which might be used as targeting exact worlds which we are interested in. In this case, the modal base would be well-formed, but attitude holder would not have an enough ordering source to pick up the best worlds.

The second case, which I think different from the first case, is *pala-* with *elyep-*. Unlike English *hope* which is not compatible with *it's hard to*, the Korean construction *palaki elyep-* is possible construction and shows a unique meaning. It has a slightly different meaning from the first case. First, the sentences contain no interrogatives in complement clause. Second, the difficulty of attitude is not related to an insufficient information or insufficient preference. Rather it has a meaning that *it is not likely to happen*. To grasp all these, I think insufficient epistemic modal base is the reason for having difficulty in attitude. The attitude holder's preference as an ordering source might be enough, but the worlds that satisfy it cannot be in the modal base (or only have the slight possibility within the modal base).

- (18) a. Mina-eykey wusung-ul pala-ki elyep-ta.
Mina-DAT win-ACC hope-NML hard-DECL
‘It is hard to hope for Mina to win’
- b. cengchi-uy mincwhuhwa-nun pala-ki elyep-ta.
politics-GEN democratization-TOP hope-NML hard-DECL
‘It is hard to hope political democratization’

This characteristic of *pala-* is again in contrast with *wenha-*, because *wenha-* is not used with *elyep-*. This

and gut-reaction related in contrast to *pala-*.

Observing mainly the data of Korean corpus, there is a distinction between hoping propositions and wanting propositions. The complement proposition of *wenha-* tends to take an attitude holder (self) as a subject and take a simple action or state predicate. Meanwhile, the complement proposition of *pala-* tends to take a subject other than an attitude holder and take a more complex predicate.

- (22) a. salam-un emi-uy pay sok-eyse na-camaca,
 human-TOP mother-GEN belly inside-LOC born-as.soon.as
 mek-ki-lul wenha-mye, al-ki-lul wenha-n-ta.
 eat-NML-ACC want-CONJ, know-NML-ACC want-PRES-DECL
 ‘As soon as one comes from the belly of a mother, she wants to eat and wants to know’
- b. nan pap mek-ki-lul wenha-pnita.
 I meal eat-NML-ACC want-DECL
 ‘I want to eat a meal’
- c. itul-un silswu-eyse paywu-ki-lul wenhay-ess-ta.
 they-TOP mistake-LOC learn-NML-ACC want-PST-DECL
 ‘They wanted to learn from mistakes’
- d. haklyekkosa-lul cekyongha-ki-lul wenha-nun tayhak-un...
 achievement.test-ACC apply-NML-ACC want-AND university-TOP
 ‘Universities that want to apply an achievement test ...’
- e. kokayk-un etteha-n ticithel swutan-ul
 customer-TOP what-ADN digital means-ACC
 sayongha-ki-lul wenha-nun-ka?
 use-NML-ACC want-PRES-INT
 ‘What digital means do customers want to use?’

- (23) a. salam-un nwukwuna coh-un salam-ulo
 man-TOP anyone good-ADN man-as
 kiektoy-ki-lul pala-n-ta.
 be.remembered-NML-ACC hope-PRES-DECL
 ‘Everyone hopes to be remembered as a good person’
- b. khun pyeng-i ani-ki-lul pala-mye
 big illness-NOM NEG.COP-NML-ACC hope-CONJ
 pyengwen-ul chac-ass-supnita.
 hospital-ACC go-PST-DECL
 ‘I went to the hospital hoping it was not a serious illness’

- c. pwuti i meymo-lul po-nun il-i
 please this memo-ACC see-ADN thing-NOM
 eps-ki-lul pala-ney
 NEG-NML-ACC hope-DECL
 'I hope you do not see this memo'
- d. kunye-uy kito-ka ilwueci-ki-lul
 she-GEN prayer-NOM be.realized-NML-ACC
 kancelhi pala-ess-ta.
 sincerely hope-PST-DECL
 'I sincerely hoped her prayer to be answered'

From the distribution of complement propositions that get along with bouletic verbs, I consider two verbs' bouletic ordering propositions are different kind ones, which results different bouletic meaning. While it is true that *wenha-* and *pala-* can have a same complement, but it is also true that there are propositions which allow only one verb. So, I want to make distinction between their bouletic ordering sources, and mark these as <boule^{epi}> and <boul^{cir}>

To explain these different kinds, my suggestion is that the different kinds of modal bases use their own kind of ordering source. Their bouletic type difference could well return to the more basic meaning-supporting concept, the modal base. So I suggest that <boule^{epi}> can come from an epistemic modal base of *pala-* and <boul^{cir}> can come from a circumstantial modal base of *wenha-*.

참고문헌

- 박재연(2004), 「한국어 양태 어미 연구」, 서울대출판부.
- Anand, Pranav & Valentine Hacquard (2013). Epistemics and attitudes. *Semantics and Pragmatics* 6:8, 1–59.
- Banerjee, Neil (2018). Trouble with Attitudes and the Future. In *Proceedings of the 35th West Coast Conference on Formal Linguistics*, ed. Wm. G. Bennett et al., 1–17. Somerville, MA: Cascadilla Proceedings Project.
- Klecha, Peter (2016). Modality and embedded temporal operators. *Semantics and Pragmatics* 9:9, 1–49.
- Portner, Paul (2009). *Modality*. Oxford surveys in semantics and pragmatics. Oxford: Oxford University Press.
- Werner, Tom (2006). Future and non-future modal sentences. *Natural Language Semantics* 14, 235–255.

인식양태 구성들의 비교 연구

-판단자와 판단증거를 중심으로-

명정희(서강대학교)

1. 서론

한국어에는 인식양태를 표현하는 우연적 구성들이 많은데 이들은 모두 인식양태를 표현한다는 공통점이 있지만 의미적 특징과 형태·통사적 특징을 좀 더 살펴보면 모두 차이를 지니고 있다. 그 차이들은 아주 다양해서 기존 논의에서도 인식양태 표현들 사이의 차이점을 밝힌 논의들은 많이 있었다. 그러나 주로 의미적 유사성을 가진 표현들 일부를 비교하는 것이 많았다. 특히 많은 논의가 이루어진 것으로, ‘-은가/나 보다’ 구문과 ‘-은가/나 싶다’ 구문을 비교한 연구나(임흥빈 1985, 차현실 1986, 엄정호 2005, 정인아 2010 등) ‘-을 것이-’와 ‘-겠-’의 의미차이에 대한 연구(이정민 1973, 이기용 1977, 서정수 1978, 고창운 1991, 염재일 2005 등)이 있다. 본 연구에서는 연구 대상의 폭을 확장하여 ‘-은가/나 보다, -은 모양이다, -은 듯하다, -은 것 같다, -을 것이다, -을 수 있다’ 6개의 인식양태 표현들의 의미를 비교하고자 한다.¹⁾

양태의 정의를 국내의 많은 논의에서 ‘화자의 명제에 대한 태도’라고 명시할 만큼 인식양태 표현에서 추측과 판단의 주체는 주로 화자라고 논의되어 왔다.²⁾

1) 이 6개의 우연적 구성은 모두 형태·통사적 구조가 다르다. 가령 ‘-은가/나 보다’의 ‘-은가, -나’는 하계체나 혹은 혼잣말에 사용되는 의문형 종결어미인데 ‘보다’와 결합하여 사용되면 인식양태 의미를 나타내고 ‘-은가/나 보다’ 두 구성 사이에서 의미 차이는 발견하기 어렵다. 한편 나머지 5개의 구성들은 모두 관형사형 어미와 의존명사, ‘-이다, 하다’ 등의 용언이 결합하였다. ‘-은 모양이다, -은 듯하다, -은 것 같다’는 ‘-을 것이다, -을 수 있다’와 달리 인식양태를 나타낼 때 관형사형 어미가 ‘-을’로 고정되지 않고 다양한 관형사형 어미가 제약 없이 올 수 있다. 가령 ‘-은 모양이다’는 ‘-을/었던/던/겠는’이 결합할 수 있고 이는 ‘-듯하다, -것 같다’도 마찬가지다. 이 글에서는 이러한 구성들을 표기할 때 표기의 편의상 ‘-은’을 대표적으로 쓰되, 다양한 관형사형 어미가 다 결합할 수 있음을 전제로 한다.

2) 김지은(1998)에서는 양태를 ‘선행하는 명제에 대한 화자의 심리적 태도’라고 정의하였고 이선웅(2001)에서는 ‘화자가 명제 내용에 영향을 미치지 않고 한 문장 내에서 표현하는 심리적, 정신적 태도’를 양태라고 정의하였다. 시정곤·김진희(2009)는 양태를 ‘화자의 의견이나 태도’라고 정의하였다.

- (1) 가. 철수는 영희를 좋아하나 봐.
 나. 철수는 공부가 재미있는 모양이야.
 다. 철수는 곧 결혼하는 듯해.
 라. 철수는 매운 음식을 못 먹는 것 같아.
 마. 철수는 엄마가 무서울 거야.
 바. 철수는 시험이 싫을 수 있어.

(1)의 예들은 현실에서 실제로 일어난 일을 기술하는 것이 아니라 판단 주체의 인식세계 안에서 가능한 일들을 표현한다. 발화 시점에 추측의 주체가 알고 있는 것들로 구성된 가능세계들 안에서 선행명제가 양립 가능하면 참이 된다. 예를 들어 (1가)의 의미를 간단하게 풀이하면, 추측/판단하는 주체가 알고 있는 바에 따르면, 철수는 영희를 좋아한다는 의미를 가진다. 여기에서 알고 있는 바에 따라 추측/판단하는 주체는 누구인가? (1가)에서 철수가 영희를 좋아한다고 판단한 주체는 화자이다. (1나~바)도 판단하는 주체는 화자이다. 하지만 ‘-은 듯하다’ 나 ‘-은 것 같다’ 는 판단 주체가 화자가 아닌 경우도 있다.

- (2) 가. 그녀는 온몸이 앙당그러지고 머리가 쭈뼛쭈뼛 들떴다. 갑자기 집 안에 자신이 호흡한 이산화탄소가 {몰려오는 듯하다/몰려오는 것 같다}. 그녀는 헉헉거렸다.
 나. 오빠의 두 눈에 붉은 핏발을 보자, 양희는 외로움과 안타까움으로 또 다리가 {꺾일 듯하다/꺾일 것 같다} 입술은 웃고 있는데 오빠의 표정은 잔뜩 일그러져 있다.
 다. 어쨌든 영애는 아버지를 변명해야 {될 듯해/될 것 같아} 마음에 내키지는 않았으나 천천히 입을 열었다.
 라. 선희는 형준이 그녀에게 필요 이상으로 {저자세인 듯해/저자세인 것 같아} 불만스러웠다.

(2가)에서 ‘이산화탄소가 몰려온다’ 고 판단하는 주체는 화자가 아니라, ‘그녀’ 이고 (2나)에서 ‘다리가 꺾인다’ 고 판단하고 느끼는 것은 ‘양희’ 이다. (2다, 라)에서도 판단자는 영애, 선희로 제3자이다. ‘-은 듯하다/-은 것 같다’ 는 판단자가 늘 화자와 일치되는 것이 아니라는 점에서 다른 인식양태 구성들과 다르다.³⁾ 인식양태 구성들이라고 알려진 양태 문장들에서 판단자인 ‘누군가’ 를 정의하는 것은 간단한 일이 아니다.

이러한 문제 의식과 관련하여 Lasersohn(2005)에서 기준 세계, 기준 시간에 기준 판단자라는 기준점을 추가하여 내포적 의미를 명세화하는 것이 필요하다고 논증하였고, Stephenson(2007)은 이 기준 판단자 개념을 받아들이고, 판단자-의존적인 해석을 명세화하기 위해 PRO_j라는 공대명사를 설정하였다. 국내 의미론에서는 양정석(2015, 2017)에서 기준 판단자를 도입하여 주관성 형용사, 양상 및 증거성을 설명한 바 있다. 이 글에서도 기본적으로 양태(양상)를 가능세계 변수에 대한 양화로 정의하고, Stephenson(2007)의 판단자(judge) 개념을 가능세계 의미론에 도입한다. 판단자 개념은 6개의 인식양태 구성들의 의미를 면밀하게 고찰하고 비교하는 데 유용한 틀이 될 것이다.

한국어의 인식양태 구성들에서 판단자는 ‘생각하다’ 술어가 쓰인 문장에 내포될 때, 다양한 시제 어미들과 결합할 때, 조건문에 실현될 때 달라지는 모습을 보인다. 각 환경에서 판단자가 실현되는 양상 및 인식양태 구성들의 의미적 차이를 살펴보기 위해 2장에서는 판단자가 다르게 실현되는 조건들을 예문들과 함께 정리하겠다. 3장에서는 판단자 외에 판단근거 역시 인식양태 구성

3) 또한 (2가~나)의 예들은 판단 주체가 느끼고 생각한 것을 비유적으로 표현한 것이고 (2다, 라) 역시 판단 주체가 생각하고 믿는 것을 표현했다. 이런 점에서 (2)의 예들은 일반적인 인식양태 문장들과 양상 의미 측면에서도 구분된다. 이 내용에 대해서는 2절에서 더 자세히 살펴보겠다.

들의 의미 차이를 가져오는 요소라고 보고 이에 대해 자세히 살펴보겠다.

2. 한국어 인식양태 구성에서의 판단자 역할

2.1. 명제태도 동사의 내포문

인식양태 표현들은 *생각하다*와 같은 명제태도 동사의 보문으로 내포되면, 누구의 지식 세계/인식 세계에서 판단이 이루어지는지가 명백하게 드러난다.

(3) 가. 비가 {올 모양이다/올 듯하다/올 것 같다/올 것이다/올 수 있다}.

나. 철수는 비가 {올 모양이라고/올 듯하다고/올 것 같다고/올 것이라고/올 수 있다고} 생각한다.

다. 영희가 {이길 모양이다/이길 듯하다/이길 것 같다/이길 것이다/이길 수 있다}

라. 철수는 영희가 {이길 모양이라고/이길 듯하다고/이길 것 같다고/이길 것이라고/이길 수 있다고} 생각한다.

(3가)와 (3다)는 각각 화자가 알고 있는 것에 따라 그의 인식세계에서 명제가 참이 된다. 하지만 (3나, 라)처럼 인식양태 문장이 명제태도 동사 구문에 내포되면 판단자는 화자가 아니라 모문의 주어가 된다. (3나)에서 비가 올 것 같다고 추측하는 주체는 화자가 아니라 모문의 주어인 철수이다. (3라)에서도 영희가 이길 것 같다고 추측하는 판단자는 모문의 주어인 철수이다. 이처럼 인식의 주체가 태도명제 구문의 내포문에서는 화자가 아닌 다른 개체로 전환된다.

이처럼 모문에서와 달리 내포문에서 화자가 아닌 다른 개체로 판단자가 전이되는 현상을 포착하기 위해서 Stephenson(2007)에서 판단자(judge) 개념을 도입하였다. 한국어의 인식양태도 가능세계 의미론에 판단자 매개변항(parameter)을 더하여 설명할 수 있다.

(4) 가. 비가 올 것 같다.

나. [-르 것 같다][비가 오-]

다. [(나)]^{w,t,j} = [-르 것 같다]^{w,t,j}([λw''].[λt''].[λj'']. [(비가 오-)]^{w''t''j''})] = 1

iff ∃<w',t',x>∈Epist^{w,t,j}: 세계 w'에서, 시간 t'에 판단자 x에게 비가 오다

(Stephenson 2007: 45 참고)

인식양태 구성 중 대표적으로 ‘-을 것 같다’를 포함하고 있는 문장의 예를 살펴보겠다. (4가)를 구조화하면 (4나)와 같고 의미를 합성하면 (4다)와 같다. 세계-시간-판단자라는 세 가지 변수에 따라 (4가)가 참이 되는데, 판단자(j)의 인식적 대안세계 중 하나에 비가 오는 세계가 있다는 것을 의미한다. (4)에서 j는 화자가 된다. 하지만 (4가)가 *생각하다* 동사에 내포되면 판단자가 전이된다.

(5) 가. 철수는 비가 올 것 같다고 생각한다.

나. [철수는 [VP생각한다 [CP [-리 것 같다고] [비가 오-]]]]

다. [(나)]^{w,t,j} = [[생각한다]^{w,t,j} ([λw''.[λt''.[λj'']. [비가 올 것 같다]^{w'',t'',j''})]]) ([철수는]^{w,t,j})

= 1 iff $\forall \langle w', t', x \rangle \in \text{Dox}_{w,t,\text{철수}} : \llbracket \text{비가 올 것 같다} \rrbracket^{w',t',x}$

= 1 iff $\forall \langle w', t', x \rangle \in \text{Dox}_{w,t,\text{철수}} : \exists \langle w'', t'', y \rangle \in \text{Epist}_{w',t',x} : \text{세계 } w'' \text{에서, 시간 } t'' \text{에 판단자 } y \text{에게 비가 오다}$

‘생각하다’가 표현하는 보문의 내용은 믿음의 대안집합으로 설명할 수 있다. 철수가 믿는 모든 믿음의 세계(대안집합)에서는 ‘비가 올 것 같다’라는 보문의 내용이 참이 되는 것이다. ‘비가 올 것 같다’는 인식양태의 문장이므로 다시 말하면, ‘비가 오다’라는 인식 대안 집합을 가지는 철수의 믿음 집합은 참이 된다. 명제태도 동사의 내포문에 인식양태가 실현되면 위의 (5다)에서 보는 것처럼 믿음 대안집합(Dox_{w,t,철수})의 주체 x는 철수가 되고, 인식 대안집합(Epist_{w',t',x})의 주체인 x와도 동일시된다. 판단자를 매개변항으로 도입하여 설명하면 명제태도 동사문에 인식양태가 내포될 때 인식양태 판단자와 모문 서술어인 생각하의 판단자가 일치하는 것을 잘 포착할 수 있다.

2.2. 판단자와 화자 일치 vs. 제3자 일치

2.1.에서는 일반적으로 인식양태 구성들은 명제태도 문장에 내포되면 판단 주체가 화자에서 주어로 전이되는 현상을 살펴보았다. 이러한 현상을 설명하기에 ‘판단자’ 개념이 유용한 것도 보았다. 이 절에서는 개별 인식양태 구성들마다 판단자가 나타나는 양식이 다른 것을 살펴보려고 한다.

‘-은 듯하다/-은 것 같다’는 다른 인식양태 구성들과 달리 내포된 환경이 아닌 환경에서도 3인칭 주어와 결합할 수 있으며 주어와 판단자가 일치하기도 한다. 다음의 예들은 모두 양태 구성 뒤에 ‘-었-’이 결합한 경우로, 추측 판단이 발화시 이전에 이루어진 경우이다. 먼저 ‘-은 듯하다, -은 것 같다’ 외의 다른 인식양태 구성들이 포함된 문장들을 살펴보자.

(6) 가. 그는 그 무렵에 이미 이것저것 작고 큰 죄가 많아 동창사회에 전혀 나타나지 못하는

처지였지만, 학교시절 비교적 가까웠던 친구가 객사한 걸 듣고, 비통함에 이끌려 불문곡직 상가로 찾아왔나 보았다.

가'. *{덕호는/나는} 그가 비통함에 이끌려 상가로 찾아왔나 보았다.

나. 그가 구체적으로 말하는 것을 보아 이번에는 뭔가 근거가 있는 모양이었다.

나'. *{덕호는/나는} 그가 뭔가 근거가 있는 모양이었다.

다. *영희는 근거가 있을 거였어.

다'. *{덕호는/나는} 영희가 근거가 있을 거였어.

라. 영희가 올 수도 있었어.

라'. *{덕호는/나는} 영희가 올 수도 있었어.

(6)의 예들은 ‘-나 보다, -은 모양이다, -을 것이다, -을 수 있다’ 구성에 ‘-었-’이 후행 결합한 예들이다. (6가, 나)는 문장의 주어가 3인칭이지만 선행 명제에 대해서 추측하는 판단자는 문장을 발화한 화자이다. 가령 (6가)에서 그가 비통함에 이끌려 상가로 찾아왔다고 추측하는 주체는 화자이지 ‘그’가 아니다. (6나)에서도 ‘그가 근거가 있다’는 명제는 화자의 지식에 따라 접근

가능한 인식 세계에서 참이 되는 것이지, 그의 지식에 의한 것이 아니다. ‘-을 것이다’ 는 (6다)에서 보듯이 ‘-었-’ 자체가 후행하지 못한다는 제약이 있다.⁴⁾ (6라)에서도 선행명제에 대한 판단자는 영희가 아닌 화자이다. 또한 (6가’ ~라’)의 예에서 보듯이 판단자를 모문의 주어로 삽입하면 모두 비문이 된다. 기본적으로 (6)에 사용된 인식양태 구성들은 판단자가 문장구조 표면에 명시적으로 나타나지 못하며 판단자는 문맥에서 드러나는 화자와 일치한다. 하지만 (7)에 쓰인 ‘-은 듯하다, -은 것 같다’ 는 (6)과 다른 양상을 보인다.

- (7) 가. 그녀는 나와 결혼을 은근히 바라는 듯했고 나는 그녀를 좋아하면서도 미래 개척과 사랑을 동시에 품어야 하는 다소 복잡한 감정을 겪고 있었다.
 가’. 덕호는 인호가 영희를 좋아하는 듯했다.
 나. 어제 밤에 비가 온 것 같았어.
 나’. 은희는 영희가 온 것 같았다.

(7가, 7나)는 (6가~6라)와 마찬가지로 문장의 주어에 3인칭이 오더라도 명제에 대해 판단하는 자는 화자와 동일시된다. 하지만 (7가’, 7나’)처럼 문두에 판단자가 ‘덕호는, 은희는’ 으로 나타날 수 있다. 이처럼 판단자가 공범주로 나타나지 않고 하나의 지시 대상으로 문두에 실현되면 판단자는 이 지시 대상들과 동일시된다.

또한 ‘-은 듯하다, -은 것 같다’ 구성은 (7)의 예처럼 판단자가 알고 있는 지식이나 정보와 관련하여 명제를 추측하는 인식양태도 표현하지만 이와는 달리 상상하거나 생각하고 느낀 것을 표현하는 경우도 있다. 양정석(2017)에서는 이러한 ‘-는 듯/은 듯/을 듯 하다’ 의 양상기반을 ‘믿음(dox)’으로 보았고 순서근거 함수로 유사함을 가진다고 설명하였다. 아래 예들은 일반적으로 직유 표현이라고 부르는 것들의 예들이다.

- (8) 가. (의사에게 환청을 호소하면서) 머릿속에서 매미가 우는 듯합니다.
 나. (방안의 냄새를 맡으면서) 장미 향기가 나는 듯하다.
 다. (쓴 음식을 맛보면서) 씹바귀를 씹는 듯하다.
 라. (팔뚝을 만져보면서) 쇠막대기를 잡고 있는 듯하다.

(양정석 2017: 168 예)

(8)의 예들은 양정석(2017)에서 여러 가지 감각을 상상하는 작용을 표현한 예들을 인용한 것이다. 이 글에서도 위와 같은 예들은 앞의 근거와 관련된 인식 양태와는 구별된다고 보고 믿음이나 생각의 의미를 나타내는 믿음 양상기반을 가진 함수로 구분하겠다. ‘-은/는/을 것 같다’ 구성 역시 ‘-은 듯하다’ 구성과 마찬가지로 믿음 양상기반을 가진 경우가 있다. 믿음 양태를 나타내는 (8)과 같은 예들도 판단자가 문장에 나타나지 않으면 자연스럽게 문맥적으로 가장 적합한 화자와 동일시된다. 하지만 아래 예들을 보면 믿음 양태의 경우에도 판단자가 명시적으로 문두에 드러날 수 있으며, 이때 판단자는 화자 외에 다른 지시 대상으로 전이된다.

4) ‘-을 거였어’ 가 결합하여 문장이 자연스럽게 성립하는 경우가 있는데 그 경우는 인식 양태의 의미가 아닌 다른 양태를 표현하는 경우이다. 가령 ‘내 마음대로 할 거였어’ 와 같은 문장에서처럼 과거에 있었던 의지를 표현한다.

- (9) 가. 은희는 마음이 찢어지는 듯했다/것 같았다.
 나. 영희는 수능 문제가 너무 어려운 듯했다/것 같았다.
 다. 은희는 A가방이 제일 예쁜 듯했다/것 같았다.
 라. 영희는 한국음식이 세상에서 제일 매운 듯했다/매운 것 같았다.

(9가)는 전체 문장의 주어인 ‘은희’가 판단자인데, 은희가 생각하고 믿는 명제들이 참이 되는 모든 가능세계에서 ‘(나의) 마음이 찢어진다’는 명제가 참이 됨을 표현한다. (9나)도 전체 문장의 주어인 ‘영희’가 양태 문장의 판단자로서 영희가 생각하고 믿는 모든 가능세계에서 ‘수능 문제가 어렵다’라는 명제가 참이 된다.⁵⁾

지금까지 살펴본 내용을 정리하면 인식양태 구성들은 모두 판단자-의존적인 해석을 나타내는 데, ‘-은 듯하다, -은 것 같다’ 류 구성들은 다른 구성과 달리 판단자가 모문에 명시적으로 드러날 수 있다. 또한 다른 4개의 우연적 구성들은 판단자가 항상 화자와 일치하지만 ‘-은 듯하다, -은 것 같다’ 구성은 모문에 드러나는 제3인칭 주어가 판단자로 정해질 수 있다. 이는 인식양태로 해석될 때, 믿음양태로 해석될 때 모두 동일하게 보이는 이 구성만의 특징이다.⁶⁾ 판단 주체가 제3자인 경우에는 인용 표현에서도 자연스럽게 사용될 수 있다.

- (10) 가. 덕호는 인호가 영희를 좋아하는 듯했다.
 나. 은희는 영희가 온 것 같았다.
 다. *덕호는 인호가 영희를 좋아하나 보았다.
 라. *은희는 영희가 온 모양이었다.

(10가, 나)는 모문의 주어와 판단자(j)가 일치하는 경우로 인용 표현에서도 자연스럽게 사용되었다. 하지만 모문의 주어가 판단자로 정해질 수 없는 ‘-나 보다, -은 모양이다’ 류 구성들은 (10다, 라)처럼 인용 표현에서도 성립하지 않는다.

‘-은 듯하다, -은 것 같다’가 항상 자유롭게 화자 외의 판단자가 정해지고 그 판단자가 모문 주어에 실현될 수 있는 것은 아니다. 이들 구성 뒤에 ‘-더-’가 결합하면 다른 인식양태 구성들과 마찬가지로 판단자가 화자로만 제약된다. (11)은 이 글의 연구 대상인 인식양태 구성 6개에 ‘-더-’가 후행 결합한 예들이다.

5) 이때 ‘어렵다’라는 주관성 형용사의 판단자도 설정될 수 있는데 *어렵다*라는 주관성을 판단하는 판단자는 모문 주어인 영희와 일치할 수도 있고 그 외의 다른 개인이 판단자로 정해질 수도 있다. 즉 (9나) 문장의 의미는 [영희는 [영희에게] 수능 문제가 어렵다고]생각한다]로 해석될 수도 있고, [영희는 [(주어 외 다른 개인)에게] 수능 문제가 어렵다고]생각한다]로 해석될 수도 있다. 한국어의 주관성 형용사의 기준 판단자와 관련해서는 양정석(2017) 논의 참고.

6) 물론 화자 외의 제3자와 판단자가 일치하는 문장들은 구어체에서는 많이 사용하지 않는다는 특징이 있다. 또한 과거 시제가 아닌 현재 시제의 문장들에서는 제3자와 판단자가 일치하는 경우가 어색하게 느껴진다. 과거 시제가 결합할 때만 제3자로 판단자가 자연스럽게 전이된다고 볼 수 있다. 중요한 것은 항상 화자와의 일치만 보이는 (6)의 양태 표현들과는 구별되는 차이점이라는 것이다.

(11) 가. 철수가 대학원 생활이 어려운가 보더라.

나. 개발 촉진회에서는 바다 낚시터나 해수욕장을 개발해서 여기를 관광지로 만들려고 하고 있는 모양이더라.

다. 최근 성격이 많이 유해지고 좋아진 듯하더라.

다'. *덕호는 인호가 영화를 좋아하는 듯하더라.

라. 어제 밤에 비가 온 것 같더라.

라'. *은희는 영화가 온 것 같더라.

마. *어제 밤에 비가 올 거더라.

바. *어제 밤에 영화가 올 수도 있더라.

판단자 매개변항에 대해 살펴보기 전에, 양태 구성마다 후행 결합하는 시제 유관어미의 제약이 다르다는 것을 정리할 필요가 있다. (11가~라)는 인식양태 구성 뒤에 '-더-'가 결합 가능하지만 (11마, 11바)의 '-을 것이-, -을 수 있-'은 '-더-'가 후행하면 비문이 된다. '-더-'는 명제에 대한 인식이 기준시보다 이전에 일어났다는 점에서 선시성을 부여한다. 가령 (11가)에서는 '철수가 대학원 생활이 어려운가 보다'라는 것에 대한 인식이 발화시 이전에 일어났음을 가리킨다. 이처럼 과거적 인식을 나타내는 '-더-'는 '-었-'과 달리 '화자의 지각 작용'이 일어났음을 전제로 한다. 따라서 '-더라'는 주관성 형용사 뒤에 쓰일 때 이인칭, 삼인칭 주어와는 어울릴 수 없는 동일 주어 제약 현상이 있다. 일반적으로 느낌은 그 주체의 내면 세계에서 이루어지는 것이므로 제3자의 인식 대상이 되기 어렵기 때문이다(한동완 1996:81-82). 최근 '-더-'는 증거성 논의에서 많이 다루어지면서 직접 증거를 가져야 한다는 제약이 제시된 바 있다(이정미 2011, 양정석 2017 등). 양정석(2017)에서는 판단자가 명제에 관한 직접 증거를 가진다는 것은 자신에 관한 서술이 아닌 사건 또는 상태 p가 참입에 대한 지각적 근거를 가지거나 판단자가 자신의 내적 감각 상태에 관한 서술p의 주체라는 뜻이라고 하였다.⁷⁾ 이 글에서는 '-더-'가 인식시의 선시성을 나타내며 동시에 선행 논의에서 밝힌 바와 같이, 직접 증거를 가진다고 본다. 이러한 '-더-'의 의미에 따르면, (11다', 11라')에서처럼 판단자가 제3자가 되는 문장은 성립하기 어렵다. '-더-'는 항상 화자를 판단자로 고려하는 맥락에서 실현되므로, '-은 듯하다, -은 것 같다'에 '-더-'가 후행하면 인식양태 구성의 판단자 역시 화자로 제한된다.

이 절에서는 판단자 개념을 도입하여 인식양태 구성들을 살펴보았는데, 판단자와 화자가 항상 일치하는 구성들과 그렇지 않은 구성들이 구분되는 것을 확인할 수 있다. 또한 판단자의 전이는 제한되는 경우가 있는데 양태 구성 뒤에 '-더-'가 결합할 경우이다. 이 절에서 살펴본 내용을 표로 정리하면 아래와 같다.

7) 양정석(2017)에서는 '-더라'를 독립된 통사 단위로 보고 단일한 어휘항목의 의미를 다음과 같이 기술한다.

[-더라]^{c*}: $w, t, j = \lambda p: p \in D_{\langle s, \langle i, \langle e, t \rangle \rangle \rangle}$ 이고, j는 p에 관한 직접 증거를 가질 것. [[SAY(s(c*), a(c*), $\forall \langle w', t', x \rangle \in \text{Epist}_{w, t', j}: p(w')(t')(x) = 1$] & t'' < now & t = now & j = s(c*)]]

[표1] 인식양태 표현들의 판단자 일치 현상

인식 양태	[1] -은가/나 보다 -은/는/을 모양이다	[2] -은/는/을 듯하다 -은/는/을 것 같다	[3] -을 것이다 -을 수 있다
현재추측	판단자 = 화자	판단자 = 화자 ⁸⁾	판단자 = 화자
‘-있-’ 후행결합	○	○	×
	판단자 = 화자	판단자 = 화자 판단자 = 제3자	
‘-더-’ 후행결합	○	○	×
	판단자 = 화자	판단자 = 화자	

2.3. 조건의 접속문에서의 인식양태

인식양태의 의미는 선행 명제가 화자나 주어(제3자)의 인식세계에서 참이 되는가와 관련이 되므로 기본적으로 판단자의 존재를 전제로 하는 표현이다. 그렇다면 인식양태는 조건의 접속문에서도 실현될 수 있는가? 이 장에서는 조건문에서 인식양태가 실현되는 양상을 판단자의 의미와 관련하여 살펴보겠다.

(12) 가. *비가 오겠으면 옷이 다 젖을 거야.

나. *비가 {올 모양이면/올 듯하면/올 것 같으면/올 거면/올 수 있으면} 옷이 다 젖을 거야.

다. *봄이 {될 모양이면/될 듯하면/될 것 같으면/될 거면/될 수 있으면} 만물이 생동한다.

위의 예들을 보면 조건문에서 선행절에는 인식양태 표현이 제한되는 것을 알 수 있다. (12가)에는 선어말어미 ‘-겠-’ 이 사용되었고 (12나, 다)에는 앞으로 일어날 일들에 대한 추측을 하는 인식양태 구성들이 사용되었다. 그런데 모든 조건 접속문에서 인식양태 표현이 제한되는 것은 아니다.

(13) 가. (너가 말한대로) 비가 {올 것 같으면/올 듯하면/올 수도 있으면} 나가지 말자.

나. {사고날 것 같으면/날 듯하면/날 수 있으면} 거리공연 하지마.

다. 곧 {이사갈 것 같으면/이사갈 듯하면/이사갈 수 있으면} 가구는 아직 사지마.

라. 김정은과의 회담이 성과가 없을 것 같으면, 아예 회담을 하지 않겠다는 강경한 발언도 했습니다.

마. 고객 연령층이 다양해지고 원하는 스타일도 많아져 매장 내 원하는 제품이 없을 것 같으면 아예 다른 브랜드를 소개시켜 주기도 한다.

(13)과 같이 조건문의 선행절에 ‘-은 듯하다/은 것 같다/올 수 있다’ 구성은 실현될 수도 있다.⁹⁾

8) 현재 추측 시점을 나타내는 문장들에서는 제3자와 판단자가 일치하는 것이 매우 어색하다. 가령, ‘철수는 영희가 은혜를 좋아하는 것 같다/듯하다’ 와 같은 문장들은 거의 사용되지 않는다.

9) ‘-을 것이다’ 는 조건문의 선행절에 결합한 예가 보이긴 하나 인식양태의 의미가 사라지고 의지를 나타내는 동적양태 의미만 남는다. 예를 들어 ‘너가 오후에 올 거면 장을 봐 줘’, ‘공기업에 갈 거면 많은 곳에 지원을 해 봐야 한다’ 와 같은 예들이 해당된다. 또한 ‘-을 수 있다’ 구성도

그렇다면 (12)와 (13)의 차이는 무엇일까? 통사 구조의 차이도 있을 수 있겠지만, 이 글에서는 판단자 매개변항을 중심으로 의미적 측면에서만 살펴보겠다. 일반적으로 조건 접속문은 (12)의 예문들처럼 판단자 매개변항을 가진 인식양태 표현이 선행절에 결합하지 못하는데, 조건문 전체가 조건문 전체를 판단하는 주체인 판단자를 가지기 때문이다.

(14) 가. 눈이 오면 추워진다.

가'. 은희는 눈이 오면 추워진다고 생각한다.

나. 눈이 오면 아마 추워질 것이다.

나'. 은희는 눈이 오면 아마 추워질 것이라고 생각한다.

(14가, 나)에서 선행절은 가정/조건을 제시하고 후행절은 선행절에서 가정된 조건에서 일어날 상황을 나타낸다. 이때 가정/조건을 세우고 그에 따른 결과를 판단하는 조건문의 주체는 보통 화자이다. (14나)는 후행절에 인식양태 표현이 결합하였는데 이때는, 판단자가 가정하고 있는 모든 가능세계에서 눈이 오는 경우, 그 중 일부 인식적 가능세계에서는 추워진다는 것이다. 이때 판단자는 화자로 동일하다. 이처럼 판단자가 화자와 동일시되는 경우, (14가', 나')처럼 *생각하다* 술어의 보문으로 내포될 경우, 제3자가 판단자가 될 수 있다.¹⁰⁾ 이는 판단자 의존의 특성을 보이는 인식양태 구문들과 같은 현상이다. 따라서 (14)와 같은 예들은 문장 전체의 판단자가 설정되며, 이 판단자는 보통 화자와 일치한다. 반면에 (13)의 예들은 선행절의 판단자가 화자일 수도 있고 화자가 아닌 다른 지시대상이 될 수도 있다. 다른 지시대상은 문장에 드러날 수도 있고, 생략되어 안 나타날 수도 있다.

(15) 가. 너가 잘 못할 것 같으면 다른 사람에게 기회를 줘라.

가'. *은희는 너가 잘 못할 것 같으면 다른 사람에게 기회를 주라고 생각한다.

나. 경기에서 질 듯하면 연습을 더 해.

나'. *은희는 경기에서 질 듯하면 연습을 더 하라고 생각한다.

(15가)에서 선행절의 판단자는 문장 주어인 '너'로 2인칭인 청자이다. (15나)는 판단자가 문장에 나타나진 않았지만 자신의 지식 및 근거에 따라 '경기에서 진다'라고 판단할 인식 주체는 청자이다. (15가, 나)에서는 선행절의 판단자와 후행절에서 발화수반력을 행하는 주체가 다르다. (15가)에서 선행절의 인식양태 판단자는 2인칭 주어이지만 후행절에서 다른 사람에게 기회를 주라고 명령하는 주체는 화자이다. (15나)에서도 후행절 명령의 주체는 화자이다. 이러한 특징들을 보면, (12, 14)의 예들과 달리 (15)의 예들에서는 선행절이 후행절의 영향을 덜 받는 것으로 보인다. 또한 (15가', 15나')는 (14가', 14나')의 예들과 달리 *생각하다* 보문으로 내포되면 비문이 된다.

조건 접속문들 중에는 시제 측면에서도 선행절이 후행절에 의존적으로 해석되는 경우도 있는 반면에, 선행절의 시제와 후행절의 시제가 독립적으로 해석되는 경우도 있다. 후자와 같은 유형의 조건 접속문에서 인식양태 구성이 선행절에 실현될 수 있다고 본다.

(16) 가. 비가 많이 오면 꽃이 필 것이다.

나. 아이가 초등학교에 들어가면 성숙해 질 거예요.

'-을 것이다'와 마찬가지로 양태의 중의적 의미를 지닌 구성이므로 능력양태의 의미로 많이 사용되지만 (13)의 예들처럼 인식양태의 예가 불가능한 것은 아니다.

10) 이러한 현상을 관찰하여 양정석(2017:179)에서는 '-으면' 조건문이 주관적 판단 구문의 하나임을 보인다고 하였다.

다. 겨울이 오면 추워질 것이다.

(17) 가. 아이가 어떻게 하는지 알았으면 이제 혼자 할 거예요.

나. 코골이가 그렇게 심했으면 병원에 갈 거예요.

다. 영화는 철수가 힘들어 하는 것 같으면 도와줄 거야.

라. (지금) 아이가 조금 아픈 듯하면 빨리 병원으로 가세요.

(16)에서는 선행절이 후행절에 의존적으로 해석되어 후행절 시제에 종속되어 해석된다. (16가)에서 ‘비가 많이 오다’의 시제는 후행절 시제에 따라 미래 시제로 해석된다. (16나, 다)도 마찬가지로 해석된다. 반면에 (17가)의 ‘아이가 어떻게 하는지 알았다’는 이미 실현된 사건으로 발화시 기준 선시성을 나타내는 과거시제가 결합하였다. (17나)도 선행절에 ‘코골이가 심했다’라고 하여 발화시 기준 과거를 나타낸다. (16, 17)의 예들은 선행절과 후행절의 의미 관계에 있어서도 차이를 보인다. (16)의 경우 선행절에 의해 논리적으로 후행절의 결과가 예상되어 논리적인 조건관계가 성립되나, (17)의 경우에는 특정한 맥락 안에서 선행절과 후행절이 관련을 맺는다.

이 절에서는 조건 접속문에서 선행절에 인식양태가 나타나지 못하는 경우도 있고 나타나는 경우도 있다는 것을 살펴보았다. 인식양태 구성은 ‘-은 듯하다, -은 것 같다, -을 수 있다’로 제한되어 나타난다. 또한 조건 접속문은 의미적으로 두 가지 종류로 구분되어야 함을 살펴보았다. 조건절 중에는 후행절에 덜 의존적으로 해석되고 맥락에 의해 인과관계를 맺는 유형이 존재한다. 이런 경우에 인식양태 구성이 선행절에 실현될 수 있으며 이때는 인식양태의 판단자가 선행절에 독립적으로 설정된다. 그 판단자는 화자와 동일시 될 수도 있지만 화자 외의 다른 지시대상과 일치될 수도 있다. 따라서 판단자 개념의 도입은 조건문에서 실현되는 인식양태 구문의 특성을 설명할 때에도 필요하다고 할 수 있다.

3. 인식양태의 판단 증거

인식양태 구성들은 보통 판단자와 화자가 동일시되는 특성을 지니지만 인식양태 구성들의 종류에 따라 판단자가 전이되는 특성을 가지기도 하였다. 이처럼 다 같은 인식양태 표현들이라고 하더라도 판단자의 일치 현상이 구성마다 다르게 나타나고 조건문과 같은 특정한 통사적 환경에서도 양태 구성에 따라 실현 여부가 달랐다. 이 절에서는 지금까지 살펴본 판단자 기준 외에 판단 증거들도 인식양태의 의미 차이를 가져온다고 보고 이에 대해 자세히 논의하고자 한다.

먼저 ‘-은가/나 보다, -은 모양이다’가 가진 판단증거의 특성부터 살펴보자. 이들은 다른 인식양태 구성들과 달리 판단자가 추측 시점에, 직접 지각한 판단증거를 가지고 있어야 한다는 것이 주요한 특징이다.¹¹⁾

11) 기존 논의에서도 ‘-은가/나 보다’ 구성은 화자가 특정한 근거에 의존해 추측할 때 쓰인다는 특징 때문에 근거와 관련된 논의가 주를 이룬다. 특히 ‘-은가/나 보다’는 주로 ‘싶다’ 구문이나 다른 인식양태 표현들과 비교되어 논의된 예가 많다. 임흥빈(1985:364)에서는 ‘싶다’ 구문과 비교하여 ‘보다’가 객관적인 판단의 의미 특성을 강하게 지닌다고 하였고 차현실(1986)에서는 ‘보다’ 구문이 명제 외의 지각적 사례를 판단의 근거로 하여 내린 화자의 불확실한 주관적 판단이라고 하였다. 엄정호(2005)에서는 ‘싶다’는 확정적이지 않은 판단이고 ‘보다’는 상황에 의존한 판단의 의미를 가진다고 보았다.

(18) 가. (빗방울이 떨어지는 소리를 들었을 경우) 비가 오나 보다.

나. (물 속에서 물고기들이 헤엄치면서 다니는 것을 보고 난 뒤에) 물이 깨끗한가 보다.

다. (땀돌 위에 신발이 놓여 있는 것을 보거나 사람의 말소리를 듣고) 사람들이 다들 집에 있나 보다. (이기갑 2006:74 예)

위의 예들은 모두 화자가 시각 또는 청각을 통해 지각할 수 있는 증거를 바탕으로 발화한 것이다. 이처럼 ‘-은가/나 보다’는 직접 지각한 증거를 바탕으로 또 다른 명제를 추측한다는 것을 알 수 있다. 그런데 이 양태 표현은 꼭 발화 상황에 직접 지각한 증거를 가질 때만 사용되는 것은 아니다. 판단자가 직접 지각한 내용 이외에 다른 사람에게 듣거나 다른 곳에서 보고 얻게 된 간접증거도 판단의 증거가 된다.

(19) 가. (아버지가 출장 가방 짐을 싸는 것을 보고) 아버지께서 내일 출장가시나 봐.

나. (결혼하는 소식을 들은 후) 개네 1년 뒤에는 결혼하나 봐.

다. (영수에게 전해 듣고) 영수는 시험 못 보면 엄마한테 혼나나 봐.

라. (이벤트 소식을 듣고) 맘스팩 신청하면 선물 주나 봐요. 관심 있으면 해보세요.

(19가)는 발화 시점에 직접 지각한 증거를 가지고 추측하고 있지만 (19나~라)는 발화 이전에 들은 근거를 가지로 추측하고 있다. 이때 전해들은 정보는 추측 내용과 직접적으로 관련이 되는 것이다. 가령 (19다)에서 화자가 영수에게 ‘시험을 못 봐서 엄마에게 혼났다’라는 정보를 전해 들었다고 가정해보자. 그런 상황에서 화자는 이 정보를 바탕으로 (19다)와 같이 추측할 수 있다.

‘-은가/나 보다’는 (20)에서처럼 상황을 가정하고 그에 대한 논리적 결과를 추측/추론할 때에는 사용할 수 없다.

(20) 가. *우리가 계속 이렇게 실적이 없다면 1년 뒤에는 해고되나 봐.

나. *엄마가 내가 산 집을 보면 좋아하시나 봐.

다. *아빠가 손주 소식을 들으면 웃으시나 보다.

(20)의 예들처럼 ‘-면’이 선행하는 조건의 접속문에서는 ‘-은가/나 보다’가 올 수 없다. (19다, 19라)와 달리 (20)에서 제약되는 이유는 (20)을 발화할 때는 판단자에게 선행명제와 직접적으로 관련된 증거가 없기 때문이다. (20)에서 화자는 선행절에서 현실에 없는 상황을 가정하고 그에 대한 논리적 추론을 후행절에서 할 것으로 기대된다. 가령 (20가)에서 판단자가 ‘실적이 없다’는 선행절과 ‘1년 뒤에 해고된다’라는 후행절의 인과 관계에 대해서 직접 보거나 들은 증거 없이 자신의 논리적 추론 과정만을 바탕으로 결론 내린다면 ‘-나 보다’와 결합할 때 문장이 기묘해진다. 하지만 문맥에 따라 판단자가 ‘실적이 없으면 1년 뒤에는 해고된다’는 내용을 이전에 직접 들은 바가 있다면, 증거를 가지고 있으므로 (20가)가 성립할 것이다. 따라서 ‘-은가/나 보다’는 (18) 예들처럼 추측 시점에 판단자 자신의 지각 경험을 통한 근거를 가지고 있거나 명제와 직접 관련된 근거를 들어서 알고 있을 때 사용할 수 있다. 이러한 인식양태의 판단증거와 관련된 특징은 ‘-은/는/을 모양이다’의 특징과 같다.

- (21) 가. 정장 자켓을 입은 남자와 가디건 입은 여자가 웃으며 이야기 나누며 지나간다. 그들은 서로 친한 모양이다.
- 나. 양치할 때마다 아픈 것이 심상치가 않아. 아무래도 충치가 생긴 모양이다.
- 다. 바람소리가 큰 것 보니까 태풍이 올 모양이야.
- 라. A: 우리집에 오면 커피밖에 대접할 게 없는데.
B: 진우 썬 커피 꽤 좋아하는 모양이야.
- 마. 그 여자 아버지 말이야, 참 재미있는 사람인 모양이야. 여기서 일하는 여자들한테까지 침을 질질 흘리고 다닌대.
- 바. 몇 번 만나서 이야기해 보니까 그동안 사귀었던 여학생이 한두 명이 아닌 모양이야, 기가 막혀서.

(21가)는 추측 시점에 화자가 시각적으로 지각한 것을 바탕으로 한 추측이고 (21나)는 통각을 직접 경험한 것이고 (21다)는 청각에 의해 증거를 획득하였다. 한편 (21라~바)의 판단 증거는 판단자의 추측 시점이나 그 이전에 전해들은 정보이다. 이처럼 ‘-은가/나 보다, -은 모양이다’ 류는 추측 시점에 판단자가 증거를 직접 지각하거나 명제와 직접 관련이 있는 내용을 전해 들어야 한다는 조건이 있는데, 이들은 판단 근거를 현재(추측 시점에) 갖고 있지 않거나 인식할 수 없다면 문장이 기묘해진다. 이는 다음 조건 접속문에 나타난 양태를 비교하여 확인할 수 있다.

- (22) 가. *영수가 미안하다고 문자한다면 오늘 안 올 모양이다.
나. *만약 구름이 낀다면 오늘 비가 올 모양이다.
다. *시끄럽게 소리지른다면 또 싸우는 모양이다.
라. 은희 이야기를 들어보니까 영수가 사귀자고 하면 사귄 모양이야.
마. 회의 결과 들었는데 실적 안 나오면 계약 파기할 모양이야.

(22가, 나, 다)는 미래 또는 현재 상황에 대한 추측을 나타내지만 추측하는 내용에 대한 증거를 현재 가진 것이 아니라 새롭게 가정하고 있으므로 문장이 성립하지 않는다. 하지만 (22라, 마)는 자연스러운데 (22라)에는 은희 이야기, (22마)에서는 회의 결과를 판단 증거로 지니고 있기 때문이다. 따라서 ‘-은가/나 보다’ 와 마찬가지로 추측 시점에 명제와 직접적으로 관련된 내용을 추측할 수 있는 판단증거를 화자가 가지고 있다면 조건 접속문의 후행절에도 ‘-은 모양이다’ 가 결합 가능하다.

한편 ‘-은가/나 보다, -은 모양이다’ 는 그 외의 양태 구성들과 달리 논리적 추론에 의한 결과를 판단할 때는 사용될 수 없기 때문에 비현실적 조건문에서 후행절에 나타날 수 없다.

- (23) 가. 엄마가 내가 산 집을 보면 {좋아하실 듯해/좋아할 것 같아/좋아할 거야/좋아할 수 있어/*좋아할 모양이야/*좋아할 건가 봐}.
- 나. 언니가 열심히 저축하면 곧 집을 {살 듯해/살 것 같아/살 거야/살 수도 있어/*살 모양이야/*살 건가 봐}
- 다. 만약 엄마를 다시 만나면 눈물이 {날 듯해/날 것 같아/날 거야/날 수도 있어/*날 모양이야/*날 건가 봐}
- 라. 만약 복학한다면 그 사람이랑 다시 {만날 듯해/만날 것 같아/만날 거야/만날 수도 있어/*만날 모양이야/*만날 건가 봐}

(23)에서는 선행절에서 현실에서 일어나지 않은 상황을 새롭게 가정한다. 이러한 가정된 상황에 따른 논리적인 결과를 후행절에서 제시한다. 선행절에서 후행절의 판단을 이끌어내는 근거는 판단자의 논리적인 추론이거나 일반적인 세상사에 대한 지식과 같이 판단자가 알고 있던 지식에 해당한다. (23)에서 자연스럽게 사용되는 인식양태 구성들은 판단자가 추측 시점에 증거를 직접 지각하지 않고 내적 증거만 가지고 있어도 제약 없이 사용된다. 지금까지 살펴본 예들을 통해 알 수 있는 결과를 정리하면, ‘-은가/나 보다, -은 모양이다’는 추측 시점에 직접 지각한 증거나 전언을 통해 알게 된 증거가 있어야 한다. 하지만 직접적인 외부적 증거가 없이, 판단자가 스스로의 상식이나 논리적 추론과 같은 내적 근거에 의해 새로운 가상 세계를 설정하고 그에 따른 결론을 추측할 때는 사용될 수 없다. 직접 지각한 증거가 필요하다는 조건은 이 두 양태 구성의 판단자가 화자로 제한되는 것을 잘 설명해준다. 또한 직접적인 외적 증거가 없는 상황에서 사용이 제약되므로, 아직 일어나지 않은 상황을 조건으로 제시하는 조건 접속문의 선행절에 사용될 수 없다.

한편 ‘-은 듯하다, -은 것 같다’는 위에서 잠깐 언급한 바와 같이, 인식양태 외에 믿음양태로도 사용될 수 있는데 그럴 경우에는 판단자의 주관적인 생각을 표현하는 것이므로 어떠한 판단 증거도 필요하지 않다.

- (24) 가. 내 생각에는 꽃 중에 장미가 제일 {예쁜 것 같아, 예쁜 듯해/*예쁠 거야, *예쁠 수 있어,
*예쁜가 봐, *예쁜 모양이야}
나. 웬지 여행가면 사랑을 {찾을 것 같아, 찾을 듯해/*찾을 거야, *찾을 수 있어, *찾나 봐,
*찾는 모양이야}

(24가)의 ‘내 생각에는’이나 (24나)의 ‘웬지’라는 표현에 의해 알 수 있듯이 (24가, 나)는 판단자가 자신의 지식이나 내적인 근거에 기대어 판단하는 것이 아니라, 자신의 주관적인 생각을 바탕으로 믿음을 표현한다. 이처럼 판단자의 믿음 세계와 공존할 수 있는 명제를 이야기할 때는 ‘-은 것 같다, -은 듯하다’만이 사용되고 다른 구성들은 사용되지 않는다.

4. 결론

이 글에서는 한국어의 다양한 인식양태 구성들을 판단자와 판단근거를 중심으로 비교해 보았다. 인식양태 구성들은 판단자를 필요로 하고 판단자에 의존적인 특성을 지닌다는 것은 기존 연구에서 논의되어 왔다. 이 연구에서는 그러한 기존 논의들을 받아들여, 판단자가 일치하는 대상이 개별 양태 구성마다 차이가 있다는 것을 살펴보았다. 판단자가 화자 또는 제3자와 일치하는 현상은 인식양태에 후행 결합하는 시제요소에 따라서도 영향을 받았다. ‘-을 것이다, -을 수 있다’는 순수한 인식 양태 의미를 나타낼 때는 ‘-었-’이나 ‘-더-’가 후행 결합하는데 제약이 있다. ‘-은 듯하다, -은 것 같다’는 ‘-었-, -더-’가 후행 결합하는데 제약이 없는데, ‘-었-’이 후행 결합할 때는 화자 외의 제3자의 판단을 표현할 수 있다. ‘-은가/나 보다, -은 모양이다’는 항상 화자와만 일치한다.

인식양태 구성들은 판단자 의존적 특성을 보이는 조건 어미 ‘-으면’ 접속문의 선행절에 결합 여부에서도 차이를 보인다. 조건 선행절이 후행절에 종속적으로 해석되지 않고 판단자가 독립적으로 설정될 수 있는 경우에는 인식양태 구성 일부가 선행절에 결합할 수 있었다. 이때 선행절에 나타나는 인식양태 구성들의 판단자는 화자 또는 제3자가 선택될 수 있다. 이러한 현상들을 통해

‘판단자’가 누구인가를 결정하는 것은 인식양태를 분석할 때 중요한 의미 변수라는 것을 알 수 있었다. 마지막으로 인식양태 개별 구성들의 차이를 가져오는 기준으로 판단증거를 세우고, 이에 따라 양태 표현들의 의미를 비교하였다. 인식양태 구성은 판단증거 측면에서 크게 두 가지로 나뉘는데, ‘-은가/나 보다, -은 모양이다’ 류와 그 외의 구성들이다. 전자는 추측 시점에 직접 지각한 증거나 전언에 의한 증거를 가지고 있어야만 하며 판단자 내적인 증거에 의한 판단에는 사용되지 않는다. 한편 나머지 구성들은 그러한 제약이 없다. 다만, ‘-은 듯하다, -은 것 같다’는 판단증거가 없이 판단자의 주관적인 믿음이나 생각, 상상을 표현할 때도 사용되는데 이는 믿음양태로 따로 구분하였다.

이 글에서는 판단자가 화자에서 제3자로 전이될 수 있는 조건으로 각 양태 구성의 특징과 시제 결합 환경 및 조건 선행절 환경을 제시하였지만, 화자가 아닌 제3자와 일치할 수 있는 원인에 대해서는 더 깊은 연구가 필요하다. 또한 조건문의 선행절에서 인식양태 구성이 나타날 수 있는 유형과 그렇지 않은 유형 사이에는 어떤 의미적·통사적 차이가 있는가에 대해서도 더 깊이 연구할 필요가 있다.

참고문헌

- 김지은(1998), 「우리말 양태용언 구문 연구」, 한국문화사.
- 시정곤·김건희(2009), “의존명사 구문의 양태적 고찰”, 『한국어학』 44, 177-212.
- 엄정호(2005), 『국어의 보문과 보문자』, 태학사.
- 염재일(2015), “‘-겠-’과 ‘-을 것’의 양태 비교 연구”, 『언어와 정보』 9-2, 한국언어정보학회, 1-22.
- 이기갑(2006), “한국어의 양태 표현”, 『담화인지언어학회 학술대회 발표논문집』, 67-83.
- 이선웅(2001), “국어의 양태 체계 확립을 위한 시론”, 『관악어문연구』 26, 317-339.
- 임홍빈(1985), “국어의 ‘통사적인’ 공범주에 대하여”, 『어학연구』 21-3, 서울대어학연구소, 331-384.
- 양정석(2015), “기준 판단자의 도입과 주관성 형용사, 양상, 증거성”, 『언어와 정보』 19-1, 119-146.
- 양정석(2017), 『주관적 판단의 문법: 주관성 형용사, 양상, 증거성』, 한국문화사.
- 차현실(1986), “양상술어의 통사와 의미·미확인 양상술어를 중심으로”, 『이화어문논집』 8, 11-34.
- 한동완(1996), 『국어의 시제 연구』, 태학사.
- Kratzer, A. (1991), Modality, von Stechow, A. et al. (eds.), *Semantics: An International Handbook of Contemporary Research*, de Gruyter, 639-650.
- Laserson, P. (2005), Context Dependence, Disagreement, and Predicates of Personal Taste, *Linguistics and Philosophy* 28, 109-152.
- Portner, P. (2009), *Modality*, Oxford University Press.
- Stephenson, T. (2007), *Towards a Theory of Subjective Meaning*, PH.D. Thesis, MIT.

Event Structure and Three Layers of Merge

Davood Kasraie Kilakjani (Seoul National University)

1. Introduction

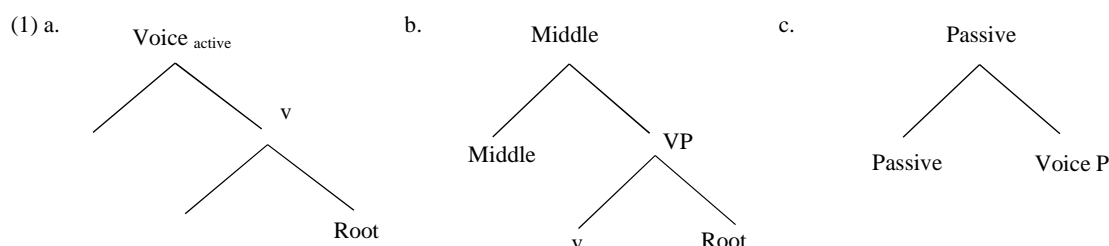
Kratzer (1996) introduces voice as a functional head, which describes the relations between a verb and its argument. Talmy (1991) makes a distinction between satellite frame languages and verb frame languages as well as introducing different kind of conflation, which are co-events to the main event. Both can be seen as a type of merge. A merge can be external (conflation) or internal (incorporation) in the sense of Chomsky (2001). This study tries to show that voice and co-events of the lexical verb can happen via conflation and incorporation. In satellite frame languages like English these merges can happen within three layers. These three layers are not equivalent of Ramchand's (2008) three subevents¹. The incorporation and conflation of the voice happen in the first layer and the second section is devoted to that. The second and third layer is related to conflation and incorporation of co-events with the lexical verb and it will be discussed in third section. The last section offers some concluding remarks and discuss some consequences of this study.

2. First Layer

Alexiadou (2006, 2014) shows that in languages like English and Greek causative/ anti-causative alternations are not derivational. She considers three voice heads:

"- Active Voice is involved in the structure of all transitive and unergative predicates across languages...Passive Voice, takes as an input a transitive structure... Middle Voice is the non-active counterpart of Kratzer's active Voice and gives rise to reflexives, passives and d. middles in Greek type languages".

Alexiadou (2014:23)

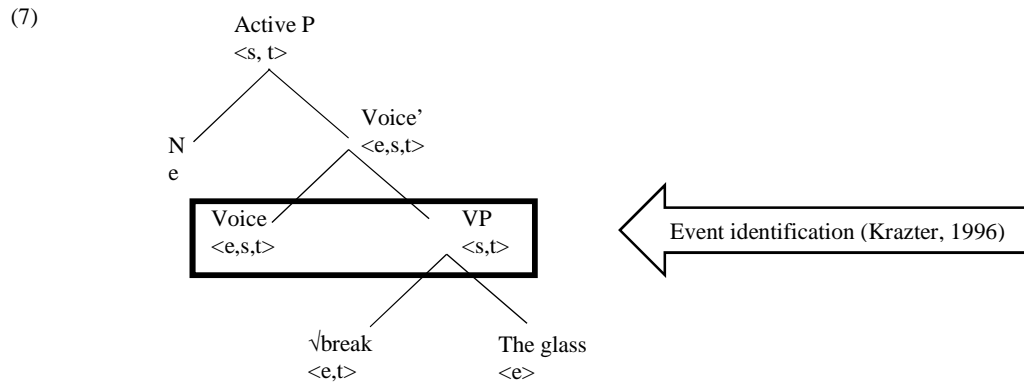


Instead of Alexiadou's middle voice, I use what one find more in the literature as inchoative. Active voices can be divided in two primitive voice heads, which are different in nature. A self-initiative voice head in which the subject is both initiator and undergoer e.g. *I ran* and a causative voice head. This distinction is necessary, as self-initiative voice head do not allow passive voice. The passive voice are built synthetically and middles are built on modals and adverbials. Thus, I assume that primitive voice heads are INCHOATIVE, SELFINITIATIVE and CAUSE:

¹ Ramchand's proposal, is independent of lexical presence of argument and resultative predicate and even for a simple event like *enter*, he considers a three subevent. Besides, he does not talk about merge or its possible effects on the event structure.

- (2) I broke the glass. (causative)
- (3) The glass broke. (inchoative)
- (4) The glass was broken. (passive)
- (5) The bread cuts easily. (middle)
- (6) I ran. (self-initiative)

When Kratzer (1996) used the term voice, she addresses the external arguments not the other voices. Perhaps that is why the derivational approach as well as non-derivational one was used to explain the whole voice phenomena.



Not all English verbs show causative/inchoative alternation. The non-alternating groups, which are listed in Levin (1996), are verbs of change of possession, verbs of cutting, verbs of contact by impact, *touch* verbs, *destroy* verbs, verbs of killing, and verbs of appearance, disappearance and occurrence, which are called verbs of external change. In addition, not all the verbs can have middle voice. *Put* verbs, *send* verbs, *Bring*, take and *carry* verbs, *drive* verbs, verb of contact etc. are among them. *Send* verbs and *carry* verbs or *hit* verbs for example cannot have either inchoative or middle voice. These verbs that do not make middle voice do so by what Levin (1988) calls lexical subordination. The resultative phrase in the (b) sentences corresponds to Levin's lexical subordination.

- (8) a. *This kind of meet pound easily.
b. This kind of meet pound thin easily.
- (9) a. *The dishes wiped easily.
b. The dishes wiped dry easily.
- (10) a. *The door kicks easily.
b. The door kicks down easily.

Levin (1988:285)

The intriguing fact about all the voice alternation is that they happen with the same verbal morphology unlike passives, which take different forms.

- (11) a. I broke the glass.
b. The glass broke.
c. The glass breaks easily.
- (12) a. I ran.
b. I ran the horses.

The data suggest that the predicate and the surface subject of an event can be understood as the summations of function². The original idea of Kratzer, which contain event identification, is also some sort of summation.

Following Kratzer, I consider *s* as the event type. The event type acts like the variable:

² (13) if $f(x)=ax$, $g(x)=bx$ then $f(x)+g(x)=ax+bx=(a+b)x$

- (14) a. Causative: $x (e_s) + (\lambda y [e' (y)] (y')) (e_s)$
 b. Inchoative/ self-initiative: $x (e_s) + [e'] (e_s)$

An example of how summation of event type works is provided below:

- (15) $I (e_s) + (\lambda y [\text{break}' (y)] (\text{the glass})) (e_s)$
 $I (e_s) + [\text{break}' (\text{the glass})] (e_s)$
 $I \text{ broke the glass } (e_s)$
 $= I \text{ broke the glass}$
 (16) $\text{the glass } (e_s) + [\text{broke}] (e_s)$
 $\text{the glass broke } (e_s)$
 $= \text{the glass broke}$

In (15), first the lexical verb take its inner argument under the causative voice conflation. Under the same event type, the subject (here the agent) adjoins to it. Also, in (16), the lexical verb and its surface subject merge under certain event type, which is inchoative.

However, not all events allow voice conflation. Some events only allow one type of voice. In (17b,c), the verbal root do not allow inchoative or middle voice alternation due to nature of the event in English which is external causation. In (18)³, however, the same verb allows inchoative voice since it refers to internal causation of the event but do not allow causative alternation. In (19), the verbal root only allow self-initiative voice head and nothing else while a verb like run do allow causative/self-initiative alternation.

- (17) a. I gave a book to John.
 b. *A book gave to John
 c. *a book gave to John easily
 (18) a. The roof gave under the heavy snow
 b. *The heavy snow gave the roof
 (19) a. He laughed
 b. *He laughed John
 c. *John laughed

In such cases, voice is a functional head that takes surface subject and the VP via composition of functions⁴:

- (21) a. Causative: $\lambda x (\lambda p [e_s (P,x)])$
 b. inchoative/ self-initiative: $\lambda x (\lambda e [e_s (e',x)])$

Here, is an example of calculation:

- (22) The roof gave...
 $(\lambda x (\lambda e [e_s (e, x)] (\text{gave})) (\text{the roof}))$
 $(\lambda x [e_s (\text{gave}, x)] (\text{the roof}))$
 $[e_s (\text{gave}, \text{the roof})]$
 $\equiv \text{the roof gave}$

If a voice conflation, do not let into causative/inchoative alternation (or something that is builds on the product inchoative voice with a merge like middles) or causative/self-initiative alternation it has its voice incorporated⁵. I will call the voice merge, the first layer.

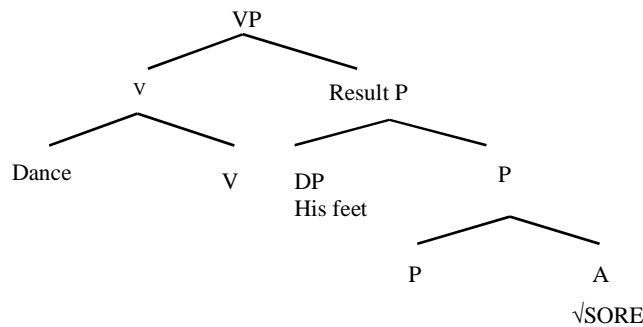
³ Sentences like *he laughed his thanks* are the result of conflation with a null light verb (Maetu 2012).

⁴ (20) $f \circ g(x) = f(g(x))$

3. Second and Third Layer

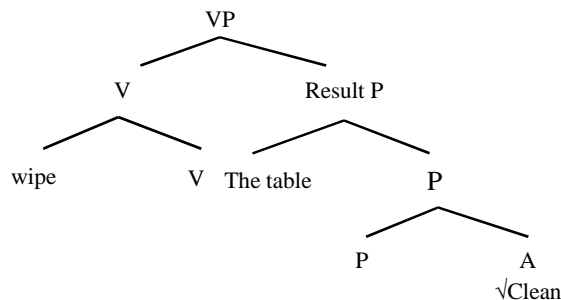
McIntyre (2004) and Mateu (2012) used an L-syntactic approach to interpret conflation of resultative predicate. Mateu argues that as results are incorporated in Japanese and Italian they do not show strong resultative. However, in English the conflation allows the strong resultatives. Here is one example from Mateu (2012):

(23) John danced his feet sore



Mateu's L-syntactic approach to result predicate claims that only Talmian satellite frame verbs can have strong resultatives and verb frame languages cannot have the strong resultatives. In L-syntactic analysis the resultative construction that involve a conflation with null light verb are strong resultatives like the one above. The DP is not the argument of the lexical verb but it is interpreted as the argument of the empty verb (light verb) inside the VP shell. What this section tries to explain is that how we should semantically treat resultative adjuncts and resultative complements. Are they all the result of conflation as Mateu claims? I will start with Mateu's (2012, 2015) L-syntactic treatment of complement resultative and adjunct resultative, which he considers the same:

"...Given the present syntactic approach to Talmy's (1991, 2000) co-event conflation pattern, which is typical of English, nothing prevents us from forming (the following)⁶ as involving conflation of the root √WIPE with a null light verb;



The analysis depicted (above) would be compatible with what Hoekstra (1988, 1992), a distinguished proponent of the Small Clause approach to resultatives, claimed: that no basic syntactic difference is to be drawn between the two types exemplified in (the following).

- a. The boy danced [_{SC/PP} his feet sore]
- b. I wiped [_{SC/PP} the table clean]

Following Hoekstra (1988, 1992) and McIntyre (2004), I will assume, for the time being, that there is no clear evidence in English for providing the two examples (above) with a different syntactic structure"

⁵ I consider verbs with missing object as causatives, which their object is muted not as self-initiative.

⁶ The author of this study added up the content of parenthesis.

However, a distinction between the two types seems necessary, as the following ungrammaticality is inevitable:

- (24) a. John Danced his feet sore.
b. *John dance his feet.
- (25) a. John hit the metal flat.
b. John hit the metal.

To show the difference between the complement resultatives and adjuncts resultatives I assume two type of meaningful light verbs in the sense of Talmy. These empty light verbs are MOTION, CAUSE, or STATE. Moreover, each of these empty light verbs can merge via conflation or incorporation under certain conditions and can have their own argument(s).

Since the voice consists the first layer of merge, I assume CAUSE or MOTION is the second layer and STATE is the third layer:

- (26) I ran up the hill into the woods.
(27) I ran <MOTION> up the hill <STATE> into the woods

In (27), MOTION is the empty light verb, which take a PP and has an implicit subject. STATE another empty verb, describes the relation between an implicit subject and a PP too. The implicit subject of each verb is the inner argument of the event (the object of lexical verb or the subject of intransitive event). Here, I abbreviate the lexical verb, cause verb, motion verb and the state verb as L, K, M, and S for convenience.

For an event like *run*, which both motion and result conflation layer is optional and the absence of none will cause ungrammaticality, the calculation of the event is summation of functions. The inner argument, which is the implicit subject of the light verb(s), is a variable, which will be taken at the end of the calculation.

- (28) a. I ran.
b. I ran up the hill.
c. I ran into the woods.
d. I ran up the hill into the woods.
- (29) $S(x) = (\lambda PP [<STATE>' (PP)] (into\ the\ woods)) (x)$
 $= [<STATE>' (into\ the\ woods)] (x)$
 $= < STATE> into\ the\ woods(x)$

$$M(x) = (\lambda PP [<MOTION>' (PP)]] (up\ the\ hill)(x)$$

$$= [<MOTION>' (up\ the\ hill)]] (x)$$

$$= < MOTION> up\ the\ hill(x)$$

$$L(x) = (x) (e_s) + ran (e_s)$$

$$= run (x) (e_s)$$

$$= run(x)$$

$$L(x) + M(x) + S(x) = (ran (x) + <MOTION>up\ the\ hill(x) + < STATE> into\ the\ woods(x)) (I)$$

$$= (ran < MOTION >up\ the\ hill <STATE >into\ the\ woods(x)) (I)$$

$$= I ran < MOTION >up\ the\ hill <STATE >into\ the\ woods$$

$$= I ran up the hill into the woods$$

In the above, MOTION and STATE are functions that take the PP with an implicit subject. The lexical verb takes its event type by voice conflation. The voiced event is the obligatory part of the summation. The meaning of the whole event is the summation of the predicates in the second and third layer.

The conflation of the third layer can happen independently of the second layer. With the difference that the second layer is empty. This helps one understand that the wiping caused cleaning the table.

- (30) a. John wiped the table clean.
 b. John wiped the table<CAUSE_{empty}> <STATE> clean
 c. John wiped the table.
 d. This table wipes clean easily.

$$(31) \quad S(y) = (\lambda AP[\langle STATE \rangle' (AP)](\text{clean})) (y) \\ = \langle STATE \rangle \text{ clean } (y)$$

$$K(y) = \langle CAUSE \rangle' \\ = \emptyset$$

$$L(y) = \text{John } (e_s) + [\text{wipe}' (y)] (e_s) \\ = \text{John wiped } (y)$$

$$L(y) + S(y) = [\text{John wiped}(y) + \langle \text{state} \rangle \text{clean } (y)](\text{The table}) \\ = [\text{John wiped } \langle \text{state} \rangle \text{clean } (y)](\text{The table}) \\ = \text{John wiped the table clean}$$

This calculation is the result of conflation of a state verb with the lexical verb. The state layer takes an adverb and an implicit object *y*. The *cause* layer is empty. The lexical verb take its syntactic subject but the calculation of the inner argument will be delayed until summation of the functions happens. At the last stage of calculation, the inner argument will be applied. As predicted by McIntyre the conflation of lexical verb root, results in a strong resultative predicate:

- (32) a. John ran the lawn flat.
 b. *John ran the lawn.
 c. John ran.

In the strong resultatives however, the AP, which belong to the STATE layer, is incorporated into the second layer. Following Ko (2015), I assume that the presence of absence of null small clause subject (here the resultative predicate) can be crucial fact to decide whether a resultative predicate is obligatory or not. In the conflation type resultatives, the STATE layer has an implicit subject. However, in incorporating type results, the STATE has an explicit subject, which is the object of the CAUSE.

$$(32) \quad S(x) = \lambda A ([\lambda y [\langle STATE \rangle' (A, y)]] (\text{flat})) (\text{the lawn}) \\ = (\lambda y [\langle STATE \rangle' (\text{flat}, y)]) (\text{the lawn}) \\ = [\langle STATE \rangle' (\text{flat}, \text{the lawn})] \\ = \text{the lawn} \langle \text{state} \rangle \text{ clean}$$

$$K(x) = (\lambda p [\langle CAUSE \rangle' (P)]) (\text{the lawn} \langle \text{state} \rangle \text{ clean})(x) \\ \langle CAUSE \rangle \text{ the lawn } \langle STATE \rangle \text{ flat}(x)$$

$$L(x) = x (e_s) + [\text{run}' (e_s)] (e_s) \\ = \text{run } (x)$$

$$L(x) + K(S(x)) = [\text{run}(x) + \langle CAUSE \rangle \text{ the lawn } \langle STATE \rangle \text{ flat}] (x) \\ = \text{John ran the lawn flat}$$

In the above example, the STATE is taken by the CAUSE via composition of functions and then the composed function sums with the product of the first layer. However, there are cases that the two layers are incorporated in into first the layer:

- (33) a. John sprinkled the turkey with salt.

- b. *John sprinkled the turkey.
- c. *The turkey sprinkled with salt.
- d. *John sprinkled.
- e. The turkey was sprinkled with salt.

(33e) shows that the predicate have an inner argument but the inner argument is not of the lexical verb (33b). Thus, the calculation would be as follows⁷:

$$\begin{aligned}
 (33) \quad S &= (\lambda PP (\lambda y [\text{<state>' (PP, y)](\text{with salt})) (\text{turkey})) \\
 &= (\lambda y [\text{<state>' (with salt, y)](\text{turkey})) \\
 &= \text{<state>' (with salt, turkey)} \\
 &= \text{turkey <state> with salt} \\
 \\
 K &= ([\lambda P[\text{<cause>' (P)}]] (\text{turkey <state > with salt})) \\
 &= \text{<cause>' (turkey <state > with salt)} \\
 &= \text{<cause> turkey <state > with salt} \\
 \\
 L &= (\lambda e (\lambda P (\lambda x [e_s(e, p, x)] (\text{sprinkle})) (\text{<cause> turkey <state > with salt})) (\text{John})) \\
 &= (\lambda P (\lambda x [e_s(\text{sprinkle}, p, x)] (\text{<cause> turkey <state > with salt})) (\text{John})) \\
 &= (\lambda x [e_s(\text{sprinkle}, \text{<cause> turkey <state > with salt}, x)](\text{John})) \\
 &= e_s(\text{John sprinkle}, \text{<cause> turkey <state > with salt}) \\
 &= \text{John sprinkle <cause> turkey <state > with salt}
 \end{aligned}$$

In the above, the STATE describes the relation between a PP and the subject of the small clause. This clause is incorporated into the CAUSE verb. The subject of SC is not the inner argument of lexical verb but the argument of the CAUSE (33b). The event type is a function that takes the lexical verb, a predicate and a subject.

However, not all the cases of merge are constructed naturally. Those that are incorporated but do not have an inner argument yell this fact by not allowing the passive voice take their surface object. Fake reflexives and some beneficiaries in object position, are made by analogy.

- (34) a. I give John a book.
- b. John was given a book.
- c. Marry made John a cake.
- d. *John was made a cake.

Although in the case of (34c) the calculation is similar to the one in (33), I believe that the resultative predicate do not have a subject as it is made by analogy:

$$\begin{aligned}
 (35) \quad S &= \lambda QP[\text{<state>' (QP)}] (\text{a cake}) \\
 &= [\text{<state>' (a cake)}] \\
 &= \text{<state > a cake} \\
 \\
 K &= ([\lambda P[\text{<cause>' (P)}]] (\text{<state > a cake})) \\
 &= [\text{<cause>' (<state > a cake)}] \\
 &= \text{<cause> John <state > a cake} \\
 \\
 L &= (\lambda e (\lambda P (\lambda x [e_s(e, p, x)] (\text{make})) (\text{<cause> John <state > a cake})) (\text{Marry}))
 \end{aligned}$$

⁷ Following Levin (1996) I assume that *spray load* verbs with variant do not have causative alternation. Thus, for this event the voice is incorporated.

=(λP (λx [e_s (make, p, x)] (<cause> John < state > with a cake)) (Marry))
 = (λx [e_s (make, <cause> John < state > a cake , x)] (Marry))
 =e_s (make, <cause> John< state > a cake, Marry)
 =Marry make <cause> John< state > a cake

4. Conclusion

In this study, I showed that incorporation and conflation of verbal root happen within three layers: the merge of voice, the merge of MOTION or CAUSE and the merge of STATE. I proposed a semantic analysis to differentiate two type of merge. For conflation, I suggest summation of functions and for incorporation, composition of functions. I also made a distinction between second and third layer merges, which are made by analogy and the cases, which are not. I showed that they are distinguishable when they are used in passive voice and in this study, I considered them as having a STATE with no subject.

I also made consider a hypothetical CAUSE layer empty for weak resultative in English. This may have some cross-linguistics consequences. As predicted Mateu (2012) verb frame languages cannot have strong resultative. This is the case for languages like Persian and Korean:

- (36) Ali Tašakkor-esh ro xandid.
 A thanks-his ACC smiled
 Ali smiled his thanks
- (37) Chelsu-nun komaum-ul wuse-cu-ess-ta.
 C TOP thanks-ACC smile-give-PAST-DEC
 Chalsu smiled his thanks

While Korean allows weak resultatives with e.g. *key* or *lo* resultatives, Persian is very conservative about weak resultative conflation and those weak resultatives may not seem acceptable for all.

- (37) ??šiše râ be hezâr tekke šekast-am.
 Glass ACC DAT thousand piece broke-1.SG
 I broke the glass into thousand pieces

I assume that cases like this are not acceptable because unlike English, Persian do not favors an empty layer of CAUSE although it does for MOTION.

Another consequence is that if we assume that there are voice heads and they conflate or incorporate with the verbal root then the derivational analysis of denominal and deadjectival verbs seems irrelevant Contrary to what Hale and Keyser (2002) claimed. If that is true the voice head merges with a noun or adjective just like verbal root and the related meaning will be understood posteriorly. For the sake of time and space, I leave these issues and further related topics for future studies.

Reference

- Alexiadou, A., Anagnostopoulou, E., & Schäfer, F. (2006). The properties of anticausatives crosslinguistically. *Phases of interpretation*, 91, 187-211.
- Alexiadou, A. (2014). Active, middle, and passive: the morpho-syntax of Voice. *Catalan journal of linguistics*, 13, 19-40.
- Chomsky, N., & Collins, C. (2001). *Beyond explanatory adequacy* (Vol. 20, pp. 1-28). MITWPL.
- Hale, K. L., & Keyser, S. J. (2002). *Prolegomenon to a theory of argument structure* (Vol. 39). MIT press.
- Hoekstra, T. (1988). Small clause results. *Lingua*, 74(2-3), 101-139.
- Kratzer, A. (1996). Severing the external argument from its verb. In *Phrase structure and the lexicon* (pp. 109-137). Springer, Dordrecht.
- Ko, H. (2015). On the Typology of Small Clauses. *생성문법연구*, 25(2), 347-375.
- Levin, B., & Rapoport, T. R. (1988). Lexical subordination. In *CLS. Papers from the General Session at the... Regional Meeting* (No. 24-1, pp. 275-289).

- Levin, B. (1993). *English verb classes and alternations: A preliminary investigation*. University of Chicago press.
- Mateu, J. (2012). Conflation and incorporation processes in resultative constructions. *Telicity, change, and state: A crosscategorical view of Event structure*, 252-278.
- Mateu, J. (2017). Two types of locative alternation. *Verb Valency Changes: Theoretical and typological perspectives*, 120, 51.
- McIntyre, A. (2004). Event paths, conflation, argument structure, and VP shells. *Linguistics*, 523-572.
- Ramchand, G. C. (2008). *Verb meaning and the lexicon: A first phase syntax* (Vol. 116). Cambridge University Press.
- Talmy, L. (1991, July). Path to realization: A typology of event conflation. In *Annual Meeting of the Berkeley Linguistics Society* (Vol. 17, No. 1, pp. 480-519).
- Talmy, L. (2000). *Toward a cognitive semantics* (Vol. 2). MIT press.

To–infinitives vs. Gerunds in Spoken and Written English: A Case Study in the ICE–GB Corpus

Gyu-Hyeong Lee/Yong-hun Lee (Hannam Univ./Chungnam Nat'l Univ.)

1. Introduction

Although it is known that both to-infinitives and gerunds are frequently used in English, their distributions and uses are different even in the native speakers' writings. Some previous studies such as Gries & Wulff (2009), Wulff & Gries (2011), Deshors (2015) and Deshors & Gries (2016) applied corpus-based and statistical approaches; and they investigated which (linguistic) factors determine the distributions and uses of to-infinitives and gerunds.

This paper extends these approaches and scrutinizes how genre/register (spoken vs. written) make different the distributions and uses of to-infinitives and gerunds. That is, the goal of this paper is to examine which (linguistic) factors determine the distributions and uses of to-infinitives and gerunds in spoken and written texts. The reason why this kind of study is necessary is that the frequency of a word or a construction is influenced by many factors including register and genre.

2. Previous Studies

2.1. Deshors (2015)

Deshors (2015) took two corpora for the analysis: one is ICE-USA for native speakers and the other is ICE-HK for non-native speakers. Then, she extracted all the sentences with to-infinitives and gerunds. After that she manually encoded six factors in Table 1 (bold faced) and NEG. She applied a multifactorial analysis to the encoded data and statistically analyzed the data.

Type	Variable (ID Tag)	Comment	Levels (ID Tag Levels)
Morphological	MatrixVerb	Verb of the Matrix Clause	
	CompVerb	Verb of the Complement Clause	
	MatrixVoice	Voice of the Main Verb	active, passive
	CompVoice	Voice of the Main Verb	active, passive
	MatrixAspect	Aspect of the Complement Verbs	neutral, perfect, progressive, progressive-perfect
	CompAspect	Aspect of the Complement Verbs	neutral, perfect, progressive, progressive-perfect
	MatrixVerbForm	Form of the Matrix Verb	finite, nonfinite
	CompIPattern	Complementation Pattern	gerund, infinitive
	Usage	Usage of Gerund/Infinitive	subj, obj, pobj, n, a, adv
Syntactic	SubjectOvert	Overt/covert Subject	covert, overt
	SubjectType	NP Type of the Subject	lexical, proper noun, pronominal, semi-pronominal
	SubjectForm	Form of the Subject	nom, acc, gen, none
	SubjectDet	Determiner of the Direct Object	definite, indefinite, no determiner
Semantic	SubjectAnimacy	Animacy of the Subject	animate, inanimate
	SubjAnimType	Subject Animacy Type	animal, flora, human, imaginary being, national, social role, absence, abstract, action, dummy it, effect, form/substance, group/measure, mental/emotional, natural entity, object/artefact, place/time, process, scholarly work, social conventions, state
	ObjectForm	Form of the Object	pp, np, cp, do (double object), pr (pronoun), no (no object)
	MatrixVerbType	Vendler's Classification	accomplishment, achievement, process, state
	CompVerbType	Vendler's Classification	accomplishment, achievement, process, state
	MatrixVerbSem	Matrix Verb Semantics	abstract, general action, action motion, action incurring transformation, action incurring movement, communication, copula, mental/cognitive/emotional, perception
	CompVerbSem	Semantics of the Complement's Lexical Verb	abstract, general action, action motion, action incurring transformation, action incurring movement, communication, copula, mental/cognitive/emotional, perception

<Table 1> Encoded Factors

Through the analysis, she observed that native speakers make syntactic decisions based on comprehensive grammatical contexts rather than single isolated semantic parameters (as previously reported). More specifically, the two types of speakers (native speakers and non-native speakers) adopted different grammatical features in the choice of a particular complement type. This implies that the two speaker populations do not share the identical knowledge of the semantic and morpho-syntactic constraints which are associated with each type of

complementation. This study also shows that the combination of cognitively oriented theoretical frameworks and empirical corpus approaches helps distinguish what motivates native and ESL speakers' syntactic choices.

2.2. Deshors & Gries (2016)

In this study, they explore verb complementation patterns of to-infinitives and gerunds both in native speakers' English (ICE-GB and ICE-USA) and three Asian non-native speakers (ICE-HK, ICE-Ind, and ICE-Sing). They also encoded a large number of variables describing the matrix verb and the complement, and they employed two random forests analyses to determine where the Asian Englishes are different from the two native speaker varieties. They found that a variety of differences between the Asian and the native Englishes. They also observed that the Asian Englishes are more similar to the American English data. They applied a two-step regression procedure to identify where non-native speakers differed in their linguistic choices from native speakers.

3. Research Methods

3.1. Corpus and Factor Encoding

As mentioned in Section 1, the frequency of a word or a construction is heavily influenced by many factors including register and genre. Because the ICE-GB corpus contains both spoken and written texts, we selected the ICE-GB corpus. In addition, since this corpus is syntactically parsed, it is easy to extract all the sentences with to-infinitives and gerunds.

From this corpus, we extracted all the sentences with to-infinitives and gerunds. Because there were too many sentences with these two constructions, we randomly selected 1,000 sentences from the extracted sentences. Then, we annotated all the linguistic factors in Table 1.

3.2. Statistical Analysis

Two types of statistical analyses can be applied to the collected data. One is a mono-factorial analysis and the other is a multi-factorial analysis, depending on how many factors/variables are involved in the analysis. In the former type of analysis, only the effects/influences of each factor/variable is statistically examined. In the latter type of analysis, however, not only the effects/influences of each factor/variable but also the interactions among the factors/variables are statistically analyzed. Though both kinds of analyses are available, multi-factorial analyses are more appropriate to the linguistic data, since many linguistic factors/variables and their interactions are involved in the linguistic phenomena. The multi-factorial analyses of linguistic data are supported by the studies in Cognitive Linguistics.

This paper also took a multi-factorial approach and used a Generalized Linear Model (GLM) with logistic regression in the statistical analysis, since it is one of the simplest and most widely-adopted analyses. the

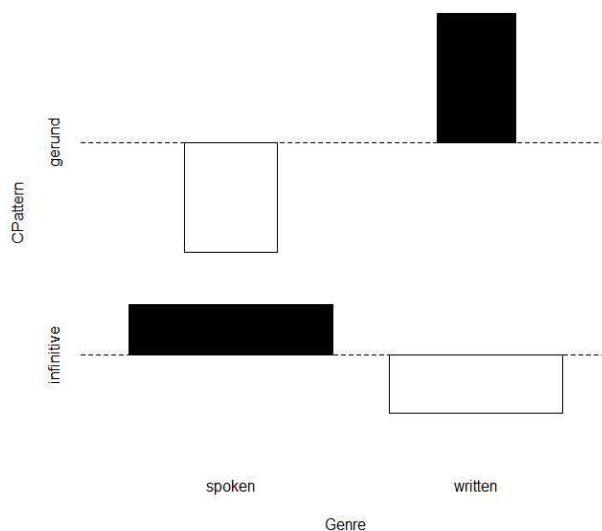
binary logistic regression is a confirmatory statistical technique that allows the analyst to identify possible correlations between the dependent and the independent factor/variables. All the statistical analyses were done with R (R Core Team, 2018).

During the analysis process, a stepwise model selection procedure was adopted as follows, which is similar to the model selection process of mixed models. First, an initial model was constructed with all of the factors and their interactions. Second, a new model was constructed where one factor or one interaction was deleted from the previous model. Third, the newly constructed model was compared with the previous model with an ANalysis Of VAriance (ANOVA). Fourth, an optimal model was chosen according to some criteria such as significance testing (with p -values) or information criteria: If a model m_1 contained a factor f or an interaction i but a model m_2 did not contain f or i , (i) when the p -value of the ANOVA test was significant ($p < .05$), it implied that the factor f or an interaction i must NOT be deleted from the model and the model m_1 was selected in this case, and (ii) when the p -value of the ANOVA was NOT significant ($p > .05$), it implied that the factor f or an interaction i can be safely deleted from the model and the model m_2 was selected in this case. The processes continued until all the factors and their interactions were scrutinized.

4. Analysis Results

4.1. Descriptive Analysis

The basic distributions of to-infinitives and gerunds are as follows.



<Figure 1> Association Plot of To-infinitives and Gerunds

As you can see, to-infinitives are more frequently used in the spoken texts.

4.2. GLM Analysis

The analysis results of GLM are as follows.

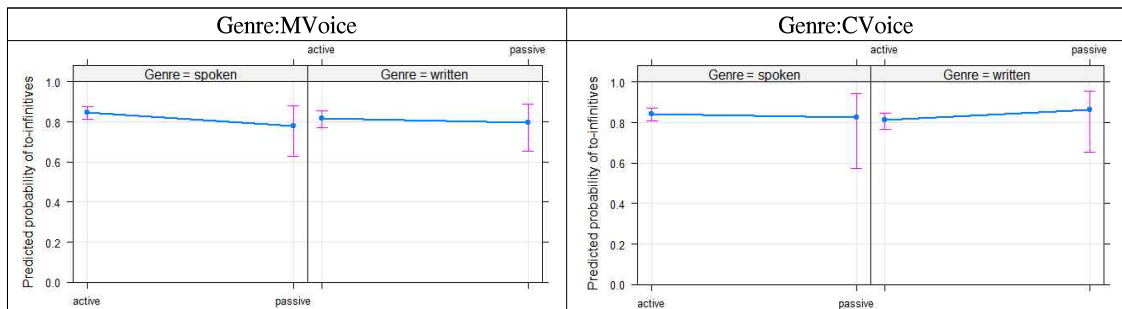
	df	Deviance	AIC	LRT	p
<none>	2595.1	2891.1			
Genre:MVoice	1	2883.5	3177.5	288.35	<.001 ***
Genre:CVoice	1	2739.3	3033.3	144.17	<.001 ***
Genre:MAspect	2	2955.6	3247.6	360.44	<.001 ***
Genre:CAspect	2	2811.4	3103.4	216.26	<.001 ***
Genre:SubjOvert	1	197.08	365.08	0.01	.961
Genre:SubjType	2	3027.7	3319.7	432.52	<.001 ***
Genre:SubjForm	2	199.55	365.55	2.48	0.290
Genre:SubjDet	2	3388.1	3680.1	792.96	<.001 ***
Genre:SubjAnim	1	2883.5	3177.5	288.35	<.001 ***
Genre:SubjAnimType	15	2955.6	3221.6	360.44	<.001 ***
Genre:ObjForm	5	3388.1	3674.1	792.96	<.001 ***
Genre:MVType	3	3748.5	4038.5	1153.40	<.001 ***
Genre:CVType	3	200.05	364.05	2.97	.396
Genre:MVSem	8	4181.1	4461.1	1585.92	<.001 ***
Genre:CVSem	9	3460.2	3738.2	865.05	<.001 ***

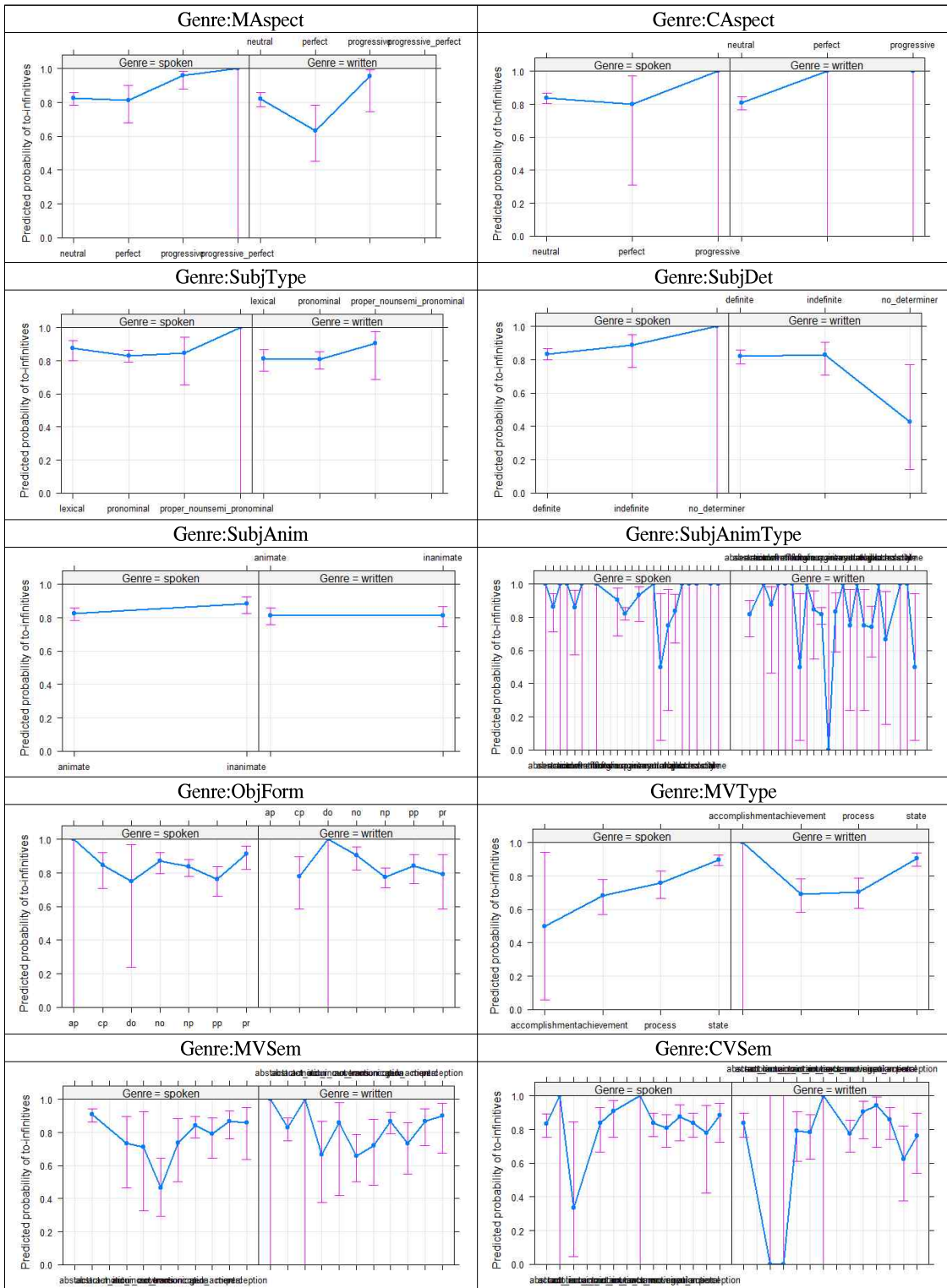
<Table 2> GLM Analysis Results

As this table illustrates, twelve linguistic factors are different between spoken and written texts.

4.3. Effect Plots

The effect plots for the (statistically) significant factors are as follows.





<Figure 2> Effect Plots for Significant Factors

These twelve factors behave differently in spoken and written texts.

5. Conclusion

As expected, the frequency of a word or a construction is influenced by many factors including register and genre, and this tendency is also observed in the distributions of to-infinitives and gerunds. To-infinitives are more frequently used in the spoken texts, and twelve factors behave differently in spoken and written texts.

References

- Deshors, S. (2015), A Multifactorial Approach to Gerundial and To-infinitival Verb-Complementation Patterns in Native and Non-native English. *English Text Construction* 8(2), 207-235.
- Deshors, S. & S. Gries (2016), Profiling Verb Complementation constructions across New Englishes: A Two-step Random Forests Analysis of *Ing* vs. *To* Complements, *International Journal of Corpus Linguistics* 21(2), 192-218.
- Gries, S. & S. Wulff (2009), Psycholinguistic and Corpus-linguistic Evidence for L2 Constructions. *Annual Review of Cognitive Linguistics* 7, 163-186.
- R Core Team (2016), R: A Language and Environment for Statistical Computing. Vienna: R Foundation for Statistical Computing.
- Wulff, S. & S. Gries (2011), Corpus-driven Methods for Assessing Accuracy in Learner Production, in P. Robinson (ed.), *Second Language Task Complexity: Researching the Cognition Hypothesis of Language Learning and Performance*. Berlin: John Benjamins, 61-87.

■ 개인 발표 4 ■

영어 I-II 교과서와 대학수학능력시험 영어 영역에 나타난 등급화된 단어의 분포

이현우 (인하대학교)

1. 서론

A. 수능 영어 영역 및 출제 범위

시행 연도별 대학수학능력시험 교육과정 및 교과서 적용 기준: 2009 개정 영어과 교육과정에 의해 제작된 2015, 2016, 2017, 2018년판 ‘영어 I’과 ‘영어 II’ 교과서임

B. 연구의 문제

‘영어 I’과 ‘영어 II’ 교과서와 절대 평가로 시행된 2018-9학년도 수능 영어 시험의 읽기 지문은 어휘 풍부성의 세 측면(즉, 어휘 정교성, 밀집성과 다양성)에 있어서 얼마나 같은지를 따져볼 필요가 있다. 따라서 본 연구는 다음의 연구 문제에 대한 답을 구하고자 한다.

첫째, ‘영어 I’과 ‘영어 II’ 교과서와 2018-9학년도 수능 영어 시험의 읽기 지문은 어휘의 정교성에 있어서 차이가 있는가?

둘째, ‘영어 I’과 ‘영어 II’ 교과서와 2018-9학년도 수능 영어 시험의 읽기 지문은 어휘의 밀집성에 있어서 차이가 있는가?

셋째, ‘영어 I’과 ‘영어 II’ 교과서와 2018-9학년도 수능 영어 시험의 읽기 지문은 어휘의 다양성에 있어서 어느 정도의 차이가 있는가?

2. 이론적 배경

A. 단어의 토큰, 타입, 레마(lemma), 군(family)

토큰은 특정한 텍스트에서 사용되는 모든 단어로 중복적으로 사용되는 단어도 다른 형태로 간주한다. 반면에, 타입은 특정한 텍스트에서 사용되는 유일무이한 단어 형태를 의미한다. 따라서 중복적으로 사용되는 동일한 단어들은 하나의 단어 타입으로 간주된다. 레마는 문법적으로 관련된 한 집단의 단어를 대표하므로, 한 단어의 모든 굴절형(즉, transform, transforms, transformed, transforming)을 포함한다. 한 단어의 군은 굴절형뿐만 아니라 규칙적인 파생으로 만들어진 형태(즉, transformer, transformation,

transformable, transformative, transformational, transformationalism)도 포함한다.

B. 어휘 정교성

- (1) The man saw the woman.
- (2) The bishop observed the actress.
- (3) The magistrate sentenced the burglar.¹⁾

비록 (1)-(3) 문장의 단어 수가 같지만, (1)의 단어들이 (2)와 (3)의 단어들보다 더 쉽고 고빈도 단어들이다. 다른 모든 조건이 같다면, (2)나 (3)의 문장을 포함하는 텍스트가 (1)의 문장을 포함하는 텍스트보다 어휘 정교성이 더 높다고 할 수 있다.

C. 어휘 밀집성

텍스트에서 차지하는 내용어(명사, 동사, 형용사, 부사)의 비율

높은 비율의 내용어를 포함하고 있는 텍스트가 높은 비율의 기능어(대명사, 조동사, 전치사, 일부 부사, 접속사, 간투사, 수사)를 포함하고 있는 텍스트보다 정보가 더 조밀하게 포장되었다.

텍스트에 다른 단어가 더 많이 사용되고 같은 단어의 반복이 적을수록 어휘 다양성이 높아진다. 이러한 어휘 다양성을 계량하는 가장 전통적인 방법은 Templin(1957)이 제안한 텍스트의 총 토큰 수 대비 단어 타입의 수의 비율, 타입-토큰 비율(type-token ratio, TTR)이다. 하지만 TTR 계량은 텍스트에 새로 추가되는 타입의 수가 점점 줄어든다는 문제를 여전히 안고 있다. 이는 더 긴 텍스트가 더 짧은 텍스트보다 일반적으로 더 낮은 TTR 값을 갖게 된다는 것을 함축한다. 이러한 문제를 해결하기 위한 하나의 방법은 Guiraud(1954)가 고안한 Guiraud 지수(GI)를 사용하는 것이다. GI는 V/\sqrt{N} 으로 표기되는 바와 같이 타입을 토큰의 제곱근으로 나눈 값이다. 이렇게 얻은 GI는 길이가 다른 텍스트의 비교에서도 유용하다.

Malvern, Richards, Chipere와 Durán(2004)이 지적했듯이, 어휘 정교성은 어휘 다양성과 밀접한 관계를 가지고 있다. 긴 텍스트에서 다양성은 다른 단어를 추가해야만 증가되는데, 다른 단어가 더 많이 추가될수록 더 고급 단어를 사용하기 때문이다.

Daller, van Hout와 Treffers-Daller(2003)은 Advanced Guiraud(AG)를 제안하였다. AG는 고급 단어 타입을 토큰의 제곱근으로 나눈 값($Vadv/\sqrt{N}$)이다. Daller(2010)은 GI가 1,000개부터 100,000개 사이의 토큰을 가진 텍스트에서 안정적인 결과를 내놓기 때문에 어휘 풍부성의 타당한 계량이지만, 빈도가 가장 높은 2,000(아마도 1,000이 더 좋음) 단어를 배제하는 것이 더 좋다고 주장한다.

3. 연구 방법 및 절차

A. 분석 대상

1) Meara와 Bell(2001)로부터 인용한 예임.

<표1> 분석 대상에서 사용된 단어

분석 대상	단어 토큰	단어 타입
YBM(신)	13,647	2,980
금성(김)	11,934	2,854
능률(이)	12,850	2,646
두산동아(윤)	14,407	3,115
미래엔(양)	12,962	2,773
비상(홍)	13,262	2,932
천재(김)	13,206	2,627
천재(이)	15,478	2,875
2018 수능	4,450	1,548
2019 수능	4,343	1,569

B. 단어의 등급화

빈도가 가장 높은 첫 번째 1,000 단어 군 목록에 있는 단어가 토큰 상 수능 영어 수능 읽기 시험의 약 75%를, 그리고 빈도가 가장 높은 2,000 단어 군 목록의 단어가 토큰 상 수능 영어 읽기 시험의 약 85%를 차지하고 있음.

2009 개정 영어과 교육과정의 기본 어휘를 기반으로 하여 단어 등급화를 다음과 같이 한다.

그룹 1: 교육과정의 기본 어휘 목록에 있는 단어로서 초등학교에서의 사용을 권장하는 단어이거나 Nation(2012)의 BNC/COCA 1st 1000 또는 2nd 1000 단어 목록에 있는 단어(1,881개)

그룹 2: 그룹 1에 속하지 않는 기본 어휘 목록에 있는 단어 또는 기본 어휘가 아닌 단어(1,024+a개)

<표2> 분석 대상에서 사용된 등급화된 단어

분석 대상	등급	단어 토큰	단어 타입
YBM(신)	1	11,848	1,941
	2	1,189	773
금성(김)	1	10,177	1,845
	2	1,350	850
능률(이)	1	11,424	1,868
	2	908	584
두산동아(윤)	1	12,210	2,010
	2	1,316	797
미래엔(양)	1	11,202	1,807
	2	1,263	784
비상(홍)	1	11,476	1,900
	2	1,185	769
천재(김)	1	11,649	1,807
	2	953	608
천재(이)	1	13,447	1,889
	2	1,211	704
2018 수능	1	3,684	1,013
	2	622	457
2019 수능	1	3,581	1,031
	2	643	471

C. 연구 가설

귀무가설 1: ‘영어 I-II’ 교과서 합본과 2018학년도 수능 영어 시험, 그리고 2019학년도 수능 영어 시험의 읽기 지문은 등급화된 단어 토큰의 빈도와 관계가 없다.

귀무가설 2: ‘영어 I-II’ 교과서 합본과 2018학년도 수능 영어 시험, 그리고 2019학년도 수능 영어 시험의 읽기 지문은 내용어와 기능어 토큰의 빈도와 관계가 없다.

D. 분석 도구 및 방법

Heatley와 Nation(2006)의 Range32H: 등급화된 단어 군, 타입, 토큰 빈도

4개의 베이스워드 목록: 앞서 언급한 두 개의 등급화된 단어 그룹과 외래어 그룹 및 고유 명사 등의 그룹이 필요하다. 베이스워드 목록은 표제어와 표제어로부터 굴절이나 교육과정의 기본 어휘 목록에서 허용되는 파생에 의해 만들어진 단어가 포함된다. 예를 들어, 베이스워드 목록에는 표제어 *nation*뿐만 아니라 *national*, *nationally*, *nationality*, *unnational*, *nationalize*, *nationalizing*, *nationalized*, *nationalizes*, *nationalization*, 등이 포함된다.

등급화된 단어 토큰과 내용어와 기능어의 단어 토큰의 빈도를 조사한 후, 이들 빈도와 2018학년도 본수능 및 6월과 9월 모평 간에 관계가 있는지를 따지기 위해 SPSS 23.0 for Windows를 활용하여 카이제곱 검정을 한다. 두 변수 간에 유의미한 관계가 있는 경우, Bonferroni의 보정법에 의한 사후 분석을 한다.

4. 연구 결과 및 논의

A. 귀무가설 1의 검증

<표3> 등급화된 단어 토큰에 대한 교차분석

	값	자유도	근사 유의확률(양측 검정)
Pearson 카이제곱	501.648 ^a	9	.000
유효 케이스 수	111,338		

a. 0셀(0.0%)은 5보다 작은 기대빈도를 가지는 셀임. 최소 기대빈도는 403.67임.

∴ 귀무가설 1 기각

Bonferroni 보정에 의한 사후 분석

1. 다중 검정 횟수 45를 곱하여 보정된 유의확률이 .000인 자료 쌍

YBM(신): 금성(김), 능률(이), 천재(김), 2018 수능, 2019 수능

금성(김): 능률(이), 두산동아(윤), 미래엔(양), 비상(홍), 천재(김), 천재(이),

2018 수능, 2019 수능

능률(이): 두산동아(윤), 미래엔(양), 비상(홍), 2018 수능, 2019 수능

두산동아(윤): 천재(김), 천재(이), 2018 수능, 2019 수능

미래엔(양): 천재(김), 천재(이), 2018 수능, 2019 수능

비상(홍): 천재(김), 2018 수능, 2019 수능

천재(김): 2018 수능, 2019 수능

천재(이): 2018 수능, 2019 수능

2. 다중 검정 횟수 45를 곱하여 보정된 유의확률이 .045인 자료 쌍
비상(홍)-천재(이)

소결: 교과서끼리 등급화된 단어 토큰의 분포가 너무 다르고, 어느 교과서도 2018학년도 또는 2019학
년도 수능 영어와 분포가 같지 않음.

B. 귀무가설 2의 검증

<표4> 분석 대상에서 사용된 기능어와 내용어

분석 대상	구분	단어 토큰
YBM(신)	기능어	6,693
	내용어	6,954
금성(김)	기능어	5,768
	내용어	6,166
능률(이)	기능어	6,372
	내용어	6,478
두산동아(윤)	기능어	6,752
	내용어	7,655
미래엔(양)	기능어	6,391
	내용어	6,571
비상(홍)	기능어	6,465
	내용어	6,797
천재(김)	기능어	6,561
	내용어	6,645
천재(이)	기능어	7,516
	내용어	7,962
2018 수능	기능어	2,097
	내용어	2,353
2019 수능	기능어	1,968
	내용어	2,375

<표5> 내용어와 기능어 토큰에 대한 교차분석

	값	자유도	근사 유의확률(양측 검정)
Pearson 카이제곱	55.224 ^a	9	.000
유효 케이스 수	116,539		

a. 0셀(0.0%)은 5보다 작은 기대빈도를 가지는 셀임. 최소 기대빈도는 2108.65임.

∴ 귀무가설 2 기각

Bonferroni 보정에 의한 사후 분석

1. 다중 검정 횟수 45를 곱하여 보정된 유의확률이 .000인 자료 쌍

YBM(신): 두산동아(윤), 2019 수능

능률(이): 두산동아(윤), 2019 수능

두산동아(윤): 미래엔(양), 천재(김)

미래엔(양): 2019 수능

비상(홍): 2019 수능

천재(김): 2019 수능

천재(이): 2019 수능

2. 다중 검정 횟수 45를 곱하여 보정된 유의확률이 .045인 자료 쌍

금성(김): 2019 수능

소결: 네 교과서 쌍에서만 기능어-내용어 토큰의 분포가 결정적인 차이를 보이고 있는 데 반해, 2019학년도 수능 영어는 금성(김)을 제외한 7개 교과서와 결정적인 차이를 보인다. 교과서보다는 2019학년도 수능 영어가 문제가 있어 보임.

C. 단어 타입-토큰의 비교

어휘 다양성을 계량하는 가장 일반적인 두 방법인 타입-토큰 비율(TTR)과 Guiraud 지수(GI)를 사용하여, 교과서와 수능 영어 시험의 어휘 다양성을 비교하고자 한다.

이들 자료의 TTR과 GI의 산출을 위해 각 자료에서 사용된 총 단어 타입 및 토큰 수 그리고 등급화된 그룹 2에 속하는 단어 타입의 수를 알아야 하는데, 이들에 관한 정보는 <표1>과 <표2>에 나와 있다. TTR과 GI 산출 공식에 따라 얻은 값은 다음과 같다.

<표6> 어휘 다양성 계량

분석 대상	TTR	GI
YBM(신)	0.218363	25.50925
금성(김)	0.239149	26.12528
능률(이)	0.205914	23.34201
두산동아(윤)	0.216214	25.95203
미래엔(양)	0.213933	24.35644
비상(홍)	0.221083	25.46006
천재(김)	0.198925	22.85991
천재(이)	0.185748	23.10896
2018 수능	0.347865	23.2055
2019 수능	0.361271	23.80828

TTR만 보면 두 수능 시험이 교과서들보다 월등히 값이 높지만, 텍스트의 길이를 보정한 GI로 보면 두 시험보다 값이 낮은 것은 천재(김)과 천재(이)뿐이다.

D. 논의

등급화된 단어의 사용에 관한 귀무가설 1의 기각과 사후 분석 결과는 영어 교과서와 2018학년도 및 2019학년도 수능 영어 시험에 사용된 자료들이 서로서로 너무나 상이하다는 것을 보여준다.

그렇다면 이들 자료들이 어휘 정교성에 있어서 어떠한 차이를 보이는지를 먼저 밝혀야 한다. 2장의 이론적 배경에서 한 텍스트의 AG 값이 클수록 고급(저빈도) 단어가 더 많이 사용되었다는 것을 뜻한다고 했다. Daller(2010)의 제안대로 고급 단어를 그룹 2 또는 그 이상의 단어로 정하고 이들 자료의 AG 값($Vadv/\sqrt{N}$)을 구해보기로 하자.

<표7> 어휘 정교성 계량

분석 대상	AG
YBM(신)	10.17802
금성(김)	12.35779
능률(이)	8.010033
두산동아(윤)	10.964
미래엔(양)	11.09347
비상(홍)	10.28996
천재(김)	8.292916
천재(이)	9.733896
2018 수능	9.324174
2019 수능	9.756994

<표6>에서 보듯이, 어휘 정교성에 있어서 능률(이)와 천재(김)은 2018-9학년도 수능 시험보다 낮고, 나머지 교과서들은 높다. 천재(이)가 어휘 정교성에 있어서 두 수능 시험과 비슷하고, YBM(신), 두산동아(윤), 비상(홍)이 조금 더 높지만, 금성(김)과 두산동아(윤)은 매우 높다고 볼 수 있다.

내용어와 기능어의 비율과 관계가 있는 귀무가설 2의 기각과 사후 분석은 몇몇 자료 간의 정보 포장이 매우 차이가 있음을 보여준다.

<표8> 어휘 밀집성 계량

분석 대상	내용어 비율(%)
YBM(신)	50.95625
금성(김)	51.6675
능률(이)	50.41245
두산동아(윤)	53.13389
미래엔(양)	50.69434
비상(홍)	51.2517
천재(김)	50.31804
천재(이)	51.44075
2018 수능	52.8764
2019 수능	54.6857

두산동아(윤)만이 2018-9학년도 수능 영어 시험과 비슷한 내용어 비율을 보이고 나머지 교과서들은 다소 낮다. 정보 제시에 있어서 2019학년도 수능 시험이 다른 자료들보다 눈에 띄게 조밀하다고 결론을 내릴 수 있을 것 같다.

결론적으로, 어휘 정교성, 밀집성, 다양성에 있어서 어느 교과서는 수능 영어 시험보다 값이 더 높게 나오고 어느 교과서는 낮게 나온다. 특히, 2018학년도 수능 시험보다 난이도가 높다는 2019학년도 수능 시험은 교과서들과 많은 차이를 보인다.

5. 결론

본 논문은 영어 I과 II 교과서 합본과 수능 영어 시험이 어휘 정교성과 밀집성에 있어서 유의미한 차이가 있는지를 카이제곱 검정을 통해 밝히고 이들 자료의 어휘 다양성을 비교하고자 했다. 카이제곱 검정 결과, 등급화된 단어 토큰의 분포(어휘 정교성)와 내용어 토큰의 비율(어휘 밀집성)에 있어서 자료들 간에 통계적으로 유의미한 차이가 있는 것으로 나타났다.

사후 분석 결과, 등급화된 단어의 사용에 있어서 교과서 간에도 차이가 많았으며, 수능 영어 시험과도 큰 차이가 있음을 알게 되었다. 어떤 교과서는 수능 시험보다도 어휘 정교성이 높았고 어떤 교과서는 낮았다. 이는 어떤 교과서를 갖고 공부하느냐에 따라서 학생들이 수능 시험 준비에 있어서 유불리가 생길 수 있다는 점을 시사하기 때문에 결코 소홀히 할 수 없는 사실이다. 따라서 향후 교과서 제작이나 수능 시험 출제에 있어서 교과서와 수능 시험 간의 어휘 연계성에 대한 치밀한 고민이 있어야 할 것으로 여겨진다.

References

- 교육과학기술부, 『영어과 교육과정. 교육과학기술부 고시 제 2011-361호 [별책 14]』, 교육과학기술부, 2011.
- 김낙복, 「대학수학능력시험 외국어(영어)영역의 코퍼스 언어학적 어휘 비교 분석」, 『영어어문교육』, 14(4), 2008, 201-20면.
- 성윤미, 「대학수학능력시험 외국어(영어) 영역에 영향을 미치는 요인들」, 『영어어문교육』, 9(2), 2004, 213-41면.
- 장경숙, 「대학수학능력시험 외국어(영어)영역 읽기 난이도 예측 모형 개발. 외국어 교육」, 11(1), 2004, 111-30면.
- 한국교육과정평가원, 「2018학년도 대학수학능력시험 6월 모의평가 채점 결과 보도자료」, 한국교육과정평가원, 서울, 2017.
- 한국교육과정평가원, 『대학수학능력시험 20년사』, 한국교육과정평가원, 서울, 2014.
- 한국교육과정평가원, 「2018학년도 대학수학능력시험 9월 모의평가 채점 결과 보도자료」, 한국교육과정평가원, 서울, 2017.
- 한국교육과정평가원, 「2018학년도 대학수학능력시험 채점 결과 보도자료」, 한국교육과정평가원, 서울, 2017.
- Abrahamsen, E. P., & Shelton, K. C., Reading comprehension in adolescents with learning disabilities: Semantic and syntactic effects, *Journal of Learning Disabilities*, 22(9), 1989, 569-72면.

- Chon, Y., & Shin, T., Item difficulty predictors of a multiple-choice reading test, *English Teaching*, 65(4), 2010, 257-82면.
- Daller, H., Milton, J., & Treffers-Daller, J., (eds.), *Modelling and assessing vocabulary knowledge*, Cambridge University Press, Cambridge, 2007.
- Daller, H., van Hout, R., & Treffers-Daller, J., Lexical richness in the spontaneous speech of bilinguals, *Applied Linguistics*, 24(2), 2003, 197-222면.
- Daller, M., Guiraud's index of lexical richness, PP presentation, University of West England, Bristol, 2010. Eprint available at <http://eprints.uwe.ac.uk/11902/>.
- Davey, B., Factors affecting the difficulty of reading comprehension items for successful and unsuccessful readers, *Journal of Experimental Education*, 56(2), 1988, 67-76면.
- Drum, P. A., Calfee, R. C., & Cook, L. K., The effects of sentence structure variables on performance in reading comprehension tests, *Reading Research Quarterly*, 16, 1981, 486-514면.
- Freedle, R., & Kostin, I., The prediction of TOEFL reading item difficulty: implications for construct validity, *Language Testing*, 10(2), 1993, 133-67면.
- Graves, M. F., Vocabulary learning and instruction, *Review of Research in Education*, 13, 1986, 49-89면.
- Guiraud, P., *Les caractères statistiques du vocabulaire: Essai de méthodologie*, Presses Universitaires de France, Paris, 1954.
- Hare, V., Rabinowitz, M., & Schieble, K., Text effects on main idea comprehension, *Reading Research Quarterly*, 24, 1989, 72-88면.
- Heatley, A., & Nation, I. S. P., *Range32H (computer software)*, Victoria University of Wellington, Wellington, New Zealand, 2006.
- Hudson, T., *Teaching second language reading*, Oxford University Press, Oxford, 2007.
- Jin, K. - A., & Park, C., The prediction of English item difficulty in college scholastic ability test, *English Teaching*, 59(1), 2004, 267-78면.
- Kieras, D. E., Thematic processes in prose, in B. Britton & J. Black (eds.), *Understanding expository text*, Erlbaum, Hillsdale, NJ, 1985, 11-64면.
- Klare, G. R., Assessing readability, *Reading Research Quarterly*, 10(1), 1974-1975, 62-102면.
- Klare, G. R., Readability, in P. D. Pearson (ed.), *Handbook of reading research*, Longman, New York, 1984, 681-744면.
- Leow, R., To simplify or not to simplify: A look at intake, *Studies in Second Language Acquisition*, 15(3), 1993, 333 - 56면.
- Malvern, D. D., Richards, B. J., Chipere, N., & Durán, P., *Lexical diversity and language development: quantification and assessment*, Palgrave Macmillan, Houndmills etc., 2004.
- Meara, P., & Bell, H., P_Lex: a simple and effective way of describing the lexical characteristics of short L2 texts, *Prospect* 16, 2001, 5-19면.
- Nation, I. S. P., *Learning vocabulary in another language*, Cambridge University Press, Cambridge, 2001.
- Nation, I. S. P., The BNC/COCA headword lists, Retrieved September 15, 2018 from <https://www.victoria.ac.nz/lals/about/staff/paul-nation#vocab-lists>, 2012.
- Ozuru, Y., Rowe, M., O'Reilly, T., & McNamara, D. S., Where's the difficulty in standardized reading tests: The passage or the question?, *Behavior Research Methods*, 40(4), 2008, 1001-15면.
- Qian, D. D., Assessing the roles of depth and breadth of vocabulary knowledge in reading comprehension, *The Canadian Modern Language Review*, 56(2), 1999, 282-308면.

- Paivio, A., *Mental representations*, Oxford University Press, New York, 1986.
- Qian, D. D., Investigating the relationship between vocabulary knowledge and academic reading performance: An assessment perspective, *Language Learning*, 52(3), 2002, 513-36면.
- Read, J., *Assessing vocabulary*, Cambridge University Press, Cambridge, 2000.
- Templin, M. C., *Certain language skills in children: Their development and inter-relationships*, University of Minnesota Press, Minneapolis, 1957.
- Ure, J., Lexical density and register differentiation, in G. E. Perren & J. L. M. Trim (eds.), *Applications of linguistics: Selected papers of the Second International Congress of Applied Linguistics*, Cambridge 1969, Cambridge University Press, Cambridge, 1971, 443-52면.
- West, M., *A general service list of English words*, Longman, Green and Co, London, 1953.
- Yano, Y., Long, M. H., & Ross, S., The effect of simplified and elaborated texts on foreign language reading comprehension, *Language Learning*, 44(2), 1994, 189-219면.

한국어 분류사구와 측량 구문의 통사-의미론적 이질성 Ramchand(2008)의 동사구 분해 이론을 적용한 해명

김용민(서울대학교)

1. 서론

한국어 화자가 어떤 사물의 개수를 세거나(counting), 관련된 물리량을 재는 것(measuring)을 언어로 표현할 때에는 관심 있는 명사와 크기를 나타낼 수사뿐만 아니라, 둘 사이를 매개하는 적절한 단어를 선택하는 것이 요구된다. 예컨대 크기가 2인 개들의 집합은 “개 두 마리”로, 4°C에서 질량 100 g에 해당하는 물 분자의 집합은 “물 일 리터”로 표현된다. 명사와 수사를 연결하는 데에 의무적으로 개입하는 이 문법 범주는 유형론적인 이유로 분류사(classifier)로 분석된다(Aikhenvald 2000).¹⁾

분류사에 의하여 명사와 수사가 결합되는 방식은 크게 명사-분류사-수사가 순서대로 결합되는 유형과 속격 표지된 수사-분류사의 연쇄(이하 ‘수량사’)가 후행하는 명사와 결합되는 유형이 있다. 최기용(2001)의 용어를 따라, 전자를 ‘비속격형’, 후자를 ‘속격형’ 구문으로 부르기로 하자. 앞서 제시한 예문인 “개 두 마리”는 비속격형에 해당하고, 이 예문에 대응되는 속격형의 구문은 “두 마리의 개”가 된다. 본고에서는 비속격형 구문만을 분석의 대상으로 삼을 것이다. 이는 비속격형 구문은 속격형 구문에 비하여 자유로운 분포를 가지고, 속격형 구문은 특히 수량화의 기능이 결여되었다는 채완(1983)의 분석을 채택한 것이다.

한국어의 경우 분류사가 단독으로 세기와 재기의 기능을 모두 수행하는 것으로 이해되고 있다.²⁾ 그러나, 많은 언어에서 세기와 재기가 의미뿐만 아니라 통사적으로도 구별된다는 보고들과, 그 차이가 실제로 범 언어적으로 관찰된다고 밝힌 Rothstein(2017)의 견해에 따르면, 이 관행을 재고할 여지가 있다.

간단히 영어에 제안된 세기와 재기의 통사론을 살펴보도록 하자. “three glasses of water”라는 표현은 영어에서 두 가지 해석을 가짐이 알려져 있다. 하나는 물이 유리잔 세 병에 각각 담겨져 있는 세기 해석과, 다른 하나는 세 잔 만큼의 물을 가리키는 재기 해석이다(Landman 2004:237). 다음의 예문에서는 문장에 주어진 맥락에 의하여 두 해석 중 하나만이 가능하다.

(1) Mary broke *three glasses of water* at once.

(2) Mary drunk *three glasses of water* at once.

(1)에서는 Mary가 물이 든 유리잔 세 병을 부수는 세기 해석만이, (2)에서는 Mary가 세 잔에 해당하는 물을 마시는 재기 해석만이 가능하다.

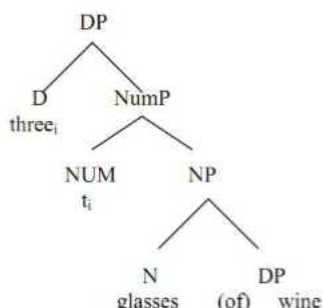
Rothstein(2011, 2016, 2017)은 (1)과 (2)에 등장하는 “three glasses of water”의 통사 구조를 아래와 같이

1) 물론 “개 둘” 혹은 “두 강아지”와 같이 분류사가 없는 구성도 관찰된다. 그러나 채완(1982, 1983)에 따르면 이 구성은 분류사 범주가 존재하지 않던 중세 국어의 특징이 화석화된 것으로 제한된 환경에서만 출현한다.

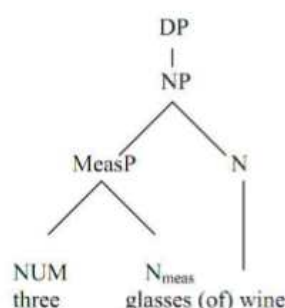
2) 한국어 분류사의 의미 체계는 다양한 범주로 분석되고 있으며, 그 범주에는 본고가 관심을 두는 세기와 재기가 일반적으로 포함된다. 대표적인 범주화 사례에 대한 소개는 최형용 외(2017)를 참조하라.

상정한다.

(3) 세기 해석 (Ramchand 2011)



(4) 재기 해석 (Ramchand 2011)



세기 해석의 구조인 (3)을 살펴보면, 불가산(uncountable) 명사 “wine”에 분류사 “glass”가 작용하여 가산성을 부여한 후 수사 “three”가 연결되는 형태이고, 재기 해석의 구조인 (4)에서는 부피를 나타내는 측량 명사(measure noun)가 수사 “three”와 먼저 결합한 후, 명사 “wine”을 수식하는 형태이다.³⁾

Rothstein(2017)은 통사론적 세기-재기 구별이 관찰되는 다른 언어에서도 (3-4)와 유사한 구조를 설정할 수 있다고 주장한다. 실제로 그는 다양한 언어에서 해당 구조를 설정할 수 있음을 밝히고 있으나, 한국어에 곧바로 적용하기는 문제가 있어 보인다. (3)의 구조는 분류사-명사 연쇄가 통사적 성분을 이루는 것을 예측하지만, 한국어는 세기와 재기에 상관없이, 비속격형 및 속격형 분류사 구문 모두에서 수량사가 확고한 통사적 구를 이루기 때문이다(채완 1983).

따라서 본고에서는 Rothstein의 기본 입장인 세기와 재기가 통사론적으로 다르게 실현될 것이라는 명제를 기본 전제로 택하되, Rothstein의 논증과는 별개로 한국어에서 해당 명제를 확인할 것이다. 그 구성은 다음과 같다. 2장에서는 세기 해석을 유도하는 분류사 구문과 재기 해석을 유도하는 구문이 현저한 통사적 차이를 보이는 예제들을 제시하고, 후자가 통사적으로 동사구 수량화 구문과 가깝게 행동하는 것을 밝힐 것이다. 3장에서는 2장에서 밝힌 세기-재기의 통사적 이질성을 설명하는 하나의 가능성으로 Ramchand(2008)의 동사구 분해를 도입한다. 4장에서는 이 가정이 설명력을 지니는 것을 보이기 위하여 세기/재기 여부, 수량사, 동사의 상이 의미적으로 상호작용하는 현상을 제시하고 특정한 사례를 분석할 것이다.

2. 세기-재기의 통사적 이질성

2.1. 비대격 및 피동 구문에서의 격 일치 여부

Sim(2004)에 따르면 비대격 및 피동 구문의 세기-재기 분류사 구문은 명사와 수량사에 각각 표지된 형태격의 일치 여부에서 차이를 보인다.

3) Rothstein(2017: 56)은 “of”가 전치사구를 투사하지 않으며, 후 통사 과정인 of-삽입(of-insertion)의 결과로 본다.

(5) 세기 해석

- 가. 얼음이 두 조각이 / *두 조각을 얼었다
- 나. 개구리가 두 마리가 / *두 마리를 잡혔다

(6) 재기 해석

- 가. 얼음이 십 인치가 / *십 인치를 얼었다
- 나. 땅이 삼 미터가 / 삼 미터를 파였다

(5)는 세기 해석을 갖는 분류사 구문, (6)은 재기 해석을 갖는 분류사 구문이 등장하고, (5가) 및 (6가)는 비대격 동사, (5나) 및 (6나)는 피동형 동사가 술어로 나타난다. (5)에서처럼 세기 해석을 함의하는 분류사 구문의 경우 명사와 수량사에 표지된 형태격은 일치할 것이 요구된다. 그러나 재기 해석을 함의하는 분류사 구문의 경우는 (6)에서처럼 명사와 수량사 간의 형태격 일치가 요구되지 않는다.

2.2. ECM 구문에서의 격 일치 여부

1절에서 Sim이 관찰한 것처럼 세기-재기는 명사와 수량사 간의 격 일치 현상에 있어서 차이를 보이는 것으로 보인다. 이 차이는 비대격/피동 구문 뿐만 아니라 예외적 격 표지(이하 ECM) 구문에서도 발견된다. 인지 동사 “생각하다”는 한국어에서 ECM을 유도하는 대표적인 동사로 알려져 있다.

(8)

- 가. 철수는 [CP *사과를 / [TP 사과가 상자 안에 세 개가 들어 있다]고] 생각한다.
- 나. 철수는 [CP 와인을 / [TP 와인이 유리병 안에 일 리터가 들어 있다]고] 생각한다.

(8)에서는 ECM 동사 “생각하다”가 이끄는 내포문의 TP는 수량사에 형태적 주격을 부여하고 있다. 한편 한국어 분류사 구문에서는 명사구 적출(extraction)이 많은 환경 하에서 가능성이 알려져 있는데, 특별히 명사구가 내포문의 TP 대신 ECM 동사 “생각하다”에 의해 대격을 인가받을 수 있는 위치인 [Spec CP] 위치로 이동하는 경우를 생각해보자.⁵⁾ 세기 분류사 구문인 (8가)에서는 그러한 이동이 차단되지만, 재기 분류사 구문인 (8나)에서는 허용된다.

2.3. 수량사의 적출 가능성

일본어는 한국어와 비슷한 비속격형 분류사 구문을 보여준다. Saito, Lin & Murasugi(2008)에 따르면 일본어 분류사 구문의 수량사는 문두로의 적출이 가능하다.

(10) (Saito, Lin and Murasugi 2008에서 차용)

- 가. Taro-wa [hon-o san saku] katta
- Taro-TOP book-ACC three CL bought
- “Taro bought three books”
- 나. san saku, Taro-wa [hon-o t] katta

4) Sim은 해당 예문의 해석이 어렵다는 것을 인정하고 있지만 별다른 표시는 하지 않았다. 의문부호는 발표자가 표시한 것이다.

5) EPP 제약은 고려하지 않는다(최기용 2009 참조).

(10가)는 전형적인 분류사 구문이 사용된 예시고, (10나)는 (10가)의 수량사 “san saku”가 문두로 적출된 것이다. 일본어와 한국어의 분류사구 구문이 구조적으로 많은 유사성을 보여줌에도 불구하고(홍사만 2008), 서셋별(2011)은 다음의 (11)과 같이 한국어에서는 (10나)에 해당하는 수량사 적출 구성이 불가능함을 지적하고, 이 차이가 일본어 명사구와 대비되는 한국어 명사구의 중요한 특징으로 보고 있다.

(11) *세 권, 철수는 책을 샀다 (서셋별 2011에서 차용)

본고의 견지에 따르면 Saito Lin & Murasugi와 서셋별은 세기 해석을 가지는 분류사 구문에 대하여 논의한 것인데, 흥미로운 것은 재기 해석을 가지는 한국어 분류사 구문의 경우 수량사의 적출이 어느 정도 용인되는 것처럼 보인다는 것이다.

(11') ?일 리터, 영화는 물을 마셨다.

(11)과 (11')의 차이는 적출된 수량사에 조사 '-이나'를 연결시켰을 때 두드러진다.

(12) ?세 권이나, 철수는 책을 샀다

(12') 일 리터나, 영화는 물을 마셨다

(12)와 (12')는 각각 (11)과 (11')의 적출된 수량사인 “세 권”, “일 리터”에 조사 '-이나'를 연결시킨 것이다. (11')에 비하여 (12')의 문법성이 상당히 상승한 것을 관찰할 수 있다. 그에 비하여 (12)의 문법성은 (11')보다는 나아 보이긴 하나, 여전히 어색해 보인다.

2.4. 동사구 수량화 구문과의 비교

놀랍게도 2.1절에서 2.3절까지 재기 분류사 구문의 통사적 행동을 (13)과 같이 동사 수량화 구문이 동일하게 보여준다.

(13)

가. 수능이 세 번이 / 세 번을 연속으로 연기됐다 (피동 구문)

나. 기상학자는 이번 폭염이 / 이번 폭염을 한 달을 갈 것이라고 예상했다 (ECM 구문)

다. 세 방울, 군인은 총에 맞았다 (수량사 적출)

(13가)와 (13나)는 피동 구문 및 ECM 구문에서 명사와 수량사 간의 격 불일치를, (13다)는 수량사의 문두 적출 가능성을 보여준다. 이는 재기 해석을 갖는 분류사 구문의 수량사가 통사적으로는 명사구보다는 동사구에 더 긴밀하게 연결되어 있다는 것을 시사한다.

3. 가정의 제안

3.1. 분류사 구문의 통사적 지위

3.1.1. 세기와 분류사구

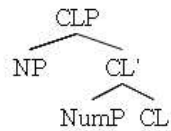
일반적으로 분류사 구문은 하나의 통사적 구를 이루는 것으로 분석된다. 다만 어떤 문법 범주가 그 구를 투사하는지에 대해서는 다양한 연구가 존재한다.⁶⁾ 그러나 2장의 자료들을 고려하면 이러한 연구 경향을 세기와 재기 해석을 갖는 각 분류사 구문 유형에 대하여 재검증할 필요가 있다.

우선 세기 해석의 분류사 구문은 통상의 연구 경향대로 통사적 구를 이루는 것으로 보인다. 기존 연구에 더하여 2.1-2.2에서 관찰한 명사-수량사간 격 일치 현상이 이 분석을 지지한다. 한편 이를 투사하는 문법 범주는 수량사 내부의 요소인 듯하다. 2.2의 관찰과 같이 명사는 비교적 이동이 자유로운 것에 비하여, 수량사는 2.3절에서처럼 적출이 불가능한 것에서 알 수 있다.

이때 이 이동은 N-이동이 아니라 NP-이동으로 보아야 한다. “연필을 철수가 두 자루 가져왔다”의 예 문처럼 적출된 명사가 동사구 밖에 안착할 수 있기 때문이다. 즉 세기 분류사 구문의 비속격형에 있어서 분류사가 매개하는 것은 수사와 명사가 아니라, 수사와 명사구라는 점이다. 분류사 구문에서 명사구의 위치는 보충어 보다는 지정어에 나타날 가능성이 높은데, 첫째는 핵을 포함하고 있는 수량사가 이미 하나의 성분을 이룬다는 점과, 지정어에 위치한 명사구가 이동에 이어서 제약을 덜 받기 때문이다.⁷⁾

한편 분류사가 물질 명사(mass noun)의 격자 구조에 원자성을 부여하여 수사와의 결합을 가능케 한다는 Chierchia(1998)과 Rothstein(2017)의 주장을 고려하면 수량사구 내부에서 분류사 구문을 투사하는 핵을 분류사로 간주하는 것이 가능할 것이다. 이 분석에 따르면 명사구가 분류사를 성분통어 함으로써 명사구가 분류사 핵에 . 남은 문제는 수사를 어떻게 처리하는 것이냐이다. 명사구가 지정어 위치에 있다는 사실을 고려하면, 수사 구가 보충어가 되는 것이 자연스럽다. 이상의 논의에서 세기의 분류사 구문은 분류사구라고 명명하는 것이 타당하다. 그 구조를 시각화하면 (14)와 같다.

(14)



3.1.2 재기와 측량 구문 및 그 문제

이제 문제가 되는 것은 재기 해석을 가지는 분류사 구문이다. 2장에서 살펴보았듯이, 이 구문은 통사적 성분을 이룬다고 보기 어렵다. 이러한 이유로 본고에서는 재기의 분류사 구문에 통사적 긴밀성을 전제하지 않는 측량 구문이라는 용어를 부여할 것이다.

2.4의 잠정적인 결론에 따라 측량 구문의 수량사가 동사구에 직접 연결된다고 간주하자. 만약 수량사가 논항 위치에 나타난다면, 이중타동사의 측량 구문은 불가능해야 한다. 그러나 이는 “철수가 어제 나무를 물을 1리터를 주었다”라는 문장으로 간단히 반박된다. 따라서 측량 수량사는 논항일 수 없고, 통사적으로는 부가어의 지위를 누릴 것이다. 그러나 이 결론 역시 문제가 있다. 만약에 측량 구문의 명사가 동사구 보충어 위치에서 대격을 인가받는 경우, 수량사는 항상 명사보다 높은 위치에 존재하게 된다. 이는 “*영희는 [v_p 한 컵을 [v 물을 마셨다]]”라는 잘못된 예측을 내놓을 뿐만 아니라, 명사 “물”과 수량사

6) 예컨대 최기용(2001)은 명사와 수량사의 격표지 양상에 따라 한정사구, 수사구의 두 가지 구조가 있음을 제안하였고, 박소영(2017)은 분류사구를 가정하고 있다.

7) 박소영(2009)는 명사와 수량사 사이에 주술관계가 성립한다는 것을 지적한다. 이는 명사구가 수량사가 투사한 분류사 구문의 지정어에 위치하면 자연스럽게 형성되는 것이다.

“한 컵” 간의 주술 관계를 파괴하기 때문에 문제가 있다.⁸⁾

3.2. Ramchand(2008)의 동사구 분해 이론

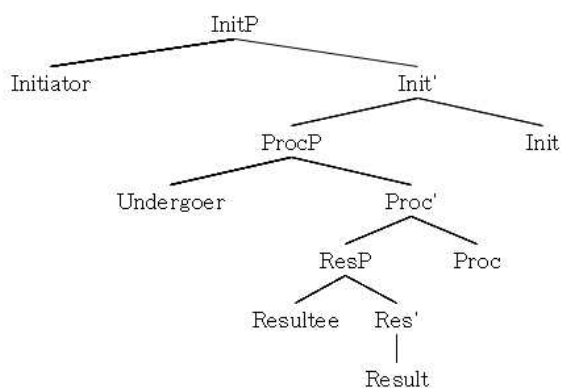
3.2.1. 소개

동사가 표상하는 사건의 의미는 촉발(CAUSE), 행동(DO), 결과(BECOME)이라는 삼부의 사건 구조(event structure)로 분해할 수 있다. 이는 의미론적인 동사구 분해로 이해할 수 있다. Ramchand(2008)는 동사의 복합적인 사건구조가 통사론에 그대로 반영되어야 한다고 강조한다. 촉발에 해당하는 InitP(Initiation Phrase), 행동에 해당하는 ProcP(Process Phrase), 결과에 해당하는 ResP(Result Phrase)가 그것이다. 이 세 구조들을 각각 하부사건(subevent)이라고 한다. 하부사건의 통사적 위계는 Init > Proc > Res 순이고, 이 순서대로 인과관계가 성립한다.

동사는 자신이 표상하는 하부사건의 핵 위치에 삽입되는데, 둘 이상의 하부사건을 지니는 의미론적으로 복합적인 동사는 그만큼 중복되어 삽입된다. 각 하부사건의 주제(theme), 즉 하부사건을 이끄는 대상 명사구는 지정어에 나타난다. Init의 지정어는 개시자(initiator), Proc은 경험자(undergoer), Res는 결과물(resultee)이다. 어떤 동사의 동작이 주어지면, 개시자는 그 동작을 시작하고, 경험자는 동작의 과정이 주는 작용을 받는다. 과정이 결과상태에 이르게 되면 결과물이 생성된다. 하부 사건 간 인과관계는 곧 주제 간의 인과관계를 함의한다.

동사가 복잡한 하부사건의 핵에 중복으로 삽입되는 것처럼, 서로 다른 하부사건의 주제가 동일할 수도 있다. 그 경우 동일한 대상 명사가 관련된 지정어 자리에 중복으로 삽입된다. 중복 삽입된 동사와 명사구는 통사적으로 가장 높은 자리에서만 발음이 나는 것으로 간주한다. 이상의 내용은 다음의 수형도로 표현된다.

(15) (Ramchand 2008 3-(1)을 한국어 통사론에 맞게 변용)



어떤 하부논항의 보충어는 다른 하부사건이 도입되는 것과 무관계할 수 있다. 그러한 보충어는 결합되어 있는 하부사건을 묘사하는 설명 요소(theme)이다. 설명 요소는 하부사건의 주제가 된 대상이 가지거나 혹은 가질 성질에 대한 정보를 제공하는데, 이는 곧 주제와 설명 요소가 서로 주술 관계를 이룸으로서 실현된다. ProcP의 경우 경로(path)라는 특별한 설명 요소를 택할 수 있다. 경로는 ProcP의 과정에

8) 박소영(2009)은 분류사구 명사-수량사간의 선택제약을 주술관계에 의한 것으로 보고 있다. 같은 논의를 측량 구문에도 다음에 쉽게 적용할 수 있다. “물 일 킬로그램/물 일 리터/*물 일 평/*물 일 뽕”.

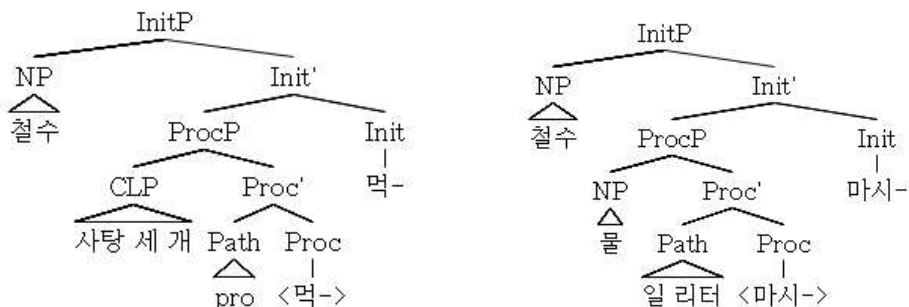
의하여 경험주가 시간에 따라 이동하는 것을 공간화한 설명 요소이다. 경로는 물리적인 수도 있지만 추상적인 수도 있다. 어떤 ProcP가 경로를 가진다면 ProcP가 표상하는 과정의 물리적/추상적으로 특징적인 구조는 그 경로에도 동등하게(homomorphic)하게 반영된다.

3.2.2. 최종 제안

3.1.2에서 문제가 된 측량 구문을 갖는 동사구의 적절한 구조를 Ramchand(2008)의 이론을 적용하면 찾을 수 있다. 측량 구문의 명사구는 하부사건을 도입하는 지정어에 삽입되지만, 수량사는 명사구를 서술할 수 있는 적절한 설명 요소가 된다. 동사에 따라서는 그러한 설명 요소가 경로가 될 수 있다. 한편 분류사구는 문제없이 구 전체가 하부사건의 지정어에 삽입될 수 있다. 두 예문 “철수가 사탕 세 개를 먹는다”와 “철수가 물 일 리터를 마신다”의 동사구 수형도를 그리면 다음과 같다.

(16) 가. “철수가 사탕 세 개를 먹는다”

나. “철수가 물 일 리터를 마신다”



이 구조의 장점은 2장의 관찰을 반영하고, 3.1.2에서 문제시되었던 측량 구문의 주술관계 문제를 해결하는 한편, 통상적인 구조의 표면형과 동일한 결과를 내놓는다는 것이다.

4. 가정의 설명력 검토

3.1. 진행상과 분류사구 및 측량 구문의 의미론적 이질성

3.2.2.의 의의는 분류사구를 포함한 동사구와 측량 구문을 포함한 동사구의 사건 구조가 다르게 형성된다는 것이다. 따라서 분류사구와 측량 구문이 사건 구조를 변형시키는 작용에 관하여 다르게 반응할 것이라는 예측을 내놓을 수 있다. 김윤신(2006, 2007)은 진행상 표지 “-고 있-”이 복합적 사건 구조를 지닌 동사의 사건 구조를 단일한 사건 구조로 사상하는 일종의 사건 함수임을 밝히고 있다.⁹⁾ 즉 진행상 표지를 결합시키는 변형은 여기서 기대하는 사건 구조를 변형시키는 작용에 해당한다. 실제로 Init과 Proc의 사건구조를 가지는 동사가 수량사로 유도되는 경로를 가질 경우 진행상에서 매우 흥미로운 현상을 보여준다. 먼저 완료상에서 다음의 예문을 보자.

9) 김윤신은 Pustejovsky(1991)의 방법론을 따르고 있지만, Saeed(2016)에 따르면 Pustejovsky와 Ramchand의 사건 분해에 대한 접근법은 서로 양립가능하다.

(18)

- 가. 영희는 국수 다섯 가닥을 뽑았다
- 나. 영희는 국수 다섯 인분을 뽑았다

(18)은 영희가 제면기로 생면 상태의 국수를 제조했던 사실을 가리킨다. (18)가는 분류사구, (18)나는 측량 구문을 사용한 것이다. 그런데 (18)에 완료상 대신에 진행상을 부여한 (18')의 해석이 흥미롭다

(18')

- 가. 영희는 (지금) 국수 다섯 가닥을 뽑고 있다
- 나. 영희는 (지금) 국수 다섯 인분을 뽑고 있다

(18'가)가 참이기 위해서는 영희가 사용하는 제면기에 구멍이 다섯 개 나 있어서 동시에 다섯 가닥의 국수가 뽑혀져 나오는 상황이어야 한다. 한편 (18'나)는 과업 목표가 국수 다섯 인분을 완성하는 것이고, 영희는 목표치에 다가가고 있는 한 시점이라는 해석을 가진다.¹⁰⁾ 이는 상을 바꾼 것만으로 동사의 양태 (manner) 정보가 다른 방식으로 함의되는 특이한 현상이다. 물론 완료상이 부여된 (18가-나)에서는 특정한 양태 해석이 두드러지지 않는다.

한편 (18), (18')에서 수사의 크기를 다섯에서 넷으로 줄여 보면 측정 구문의 진행상이 특이한 현상을 보인다. (19)와 같이 일반적으로 수량사의 크기가 줄어들면 줄어드는 방향으로 함의 관계가 생긴다.

(19)

- 가. 영희는 국수 다섯 가닥을 뽑았다 ⇒ 영희는 국수 네 가닥을 뽑았다
 - 나. 영희는 국수 다섯 인분을 뽑았다 ⇒ 영희는 국수 네 인분을 뽑았다
- 하지만 진행상에서는 분류사구와 측량 구문이 다르게 반응한다.

(19')

- 가. 영희는 국수 다섯 가닥을 뽑고 있다 ? ⇒ 영희는 국수 네 가닥을 뽑고 있다¹¹⁾
- 나. 영희는 국수 다섯 인분을 뽑고 있다 * ⇒ 영희는 국수 네 인분을 뽑고 있다¹²⁾

(19')가는 문제없어 보이는 추론이나, (19'나)는 (18'나)에서 확인하였던 과업 목표 양태 의미 때문에 잘못된 추론으로 보인다. (18나')는 과업 종료 시점이 다섯 인분을 전부 뽑아낸 순간임을 나타내기 때문이다.

3.2. (19')의 분석 가능성

10) 두 예문은 현재까지 뽑은 국수가 다섯 가닥/다섯 인분이라는 축적(cumulative) 해석도 가지나, 표준적인 진행상의 해석으로 인정하기 어렵다. 시간부사 “지금”의 존재는 이 해석을 차단한다.

11) 발표자는 정상적인 추론으로 느끼지만, (18'가)를 ‘정확히’ 다섯 가닥을 뽑는 것으로 해석할 가능성을 배제할 수가 없어서 ‘?’ 표시를 하였다.

12) “영희는 찰흙 500g을 주무르고 있다 ⇒ 영희는 찰흙 400g을 주무르고 있다”는 추론은 타당해 보이지만, “찰흙”이 “주무르다”가 표상하는 과정에 의하여 어떤 경로를 이동하는 것이 아니므로, 측정 수량사 500g은 경로가 아니라 비공간적 설명 요소로 보아야 한다.

3.2.1 관찰

(19')에서 보여진 이상 현상에 대한 하나의 가능한 분석을 제안함으로써 본고를 마치려고 한다. 이 분석은 두 가지 관찰에 근거한 것이다. 먼저 (18)에서 나타나듯 감소하는 수사가 유도하는 함의 관계는 수사의 본질적인 성질이 아니라는 점이다. Rothstein(2017:18-25)에 따르면 수사는 $\langle e, t \rangle$ 타입을 가지는 술어로 분석되는 것이 타당하다. 어떤 자연수 K 가 주어졌을 때, 그에 해당하는 수사를 k 이라 하면, k 의 의미는 대상의 크기가 K 인지를 확인하는 것이다. 즉 그 논리식은 단순히 $[[k]] = \lambda x. |x|=K$ 로 표현된다. 이 식은 (18)의 함의 관계를 유도하지 않는다. 감소하는 수사에 의해 함의 관계가 유도되려면, 그 수사는 “적어도 k ”라는 해석을 가져야 하는데, 이것은 수사가 양화사로 타입이 전환되어야 가능한 것이다. 수사 k 가 적당한 명사 N 과 결합하여 양화사로 전환되면 $\langle \langle e, t \rangle, t \rangle$ 타입인 $\lambda P. \exists x [P(x) \wedge N(x) \wedge |x|=K]$ 가 된다. 여기서 존재 양화사 \exists 가 “적어도 k ”라는 해석을 유도한다.

한편 양화사가 통사론에서 기능하기 위해서는 적당한 위치에서 성분통어에 의한 작용역(scope)이 형성되어야 하는데, 보통 기저부에서 양화사는 작용역을 형성하기에 낮은 위치에 자리 잡고 있다. 그래서 논리 형식(logical form)에서의 비현시 이동(covert movement)이 요구된다. 따라서 (19)와 (19'가)의 함의 관계는 수사가 포함된 분류사구와 측정 수량사가 양화사로의 타입 전환을 겪은 후 통사론적으로 상위의 지점으로 비현시 이동을 한 뒤 적절한 작용역을 형성하였기 때문이다. 또한 (19나')에서 함의 관계가 성립하지 않는 이유는 측정 수량사가 양화사로 전용되지 못하고 기저 위치에서 그대로 해석되기 때문일 것으로 짐작된다.

다른 한 관찰은 진행상이 ProcP보다 하위 요소인 ResP이나 경로와 제대로 호응하지 못하는 현상으로, 불완료상 모순(imperfective paradox)로 잘 알려져 있다. 전통적인 동사 어휘상(aktionsart) 중 진행상으로 사용될 수 있는 것은 행위(activity), 완성(accomplishment), 달성(achievement) 동사이다. (20)에서 볼 수 있는 것처럼 행위동사는 진행상이 완료상을 함의하지만, 완성과 달성 동사는 그렇지 않다.

(20)

- 가. 흰둥이는 바닥에서 자고 있다 \Rightarrow 흰둥이는 바닥에서 잤다 (행위동사)
- 나. 영화는 벽을 칠하고 있다 $\ast \Rightarrow$ 영화는 벽을 칠했다 (완성동사)
- 다. 철수는 집에 도착하고 있다 $\ast \Rightarrow$ 철수는 집에 도착했다 (달성동사)

한편 Ramchand는 완성동사와 달성동사들의 하부사건에서 특징적인 구조를 각각 ProcP와 경로, ResP의 존재로 꼽고 있다. 이 점을 생각하면, (20)에서 나타나는 불완료상 모순은 진행상 표지 “-고 있-”이 경로와 ResP의 해석을 방해하는 결과로 보인다.

3.2.2 제안과 분석

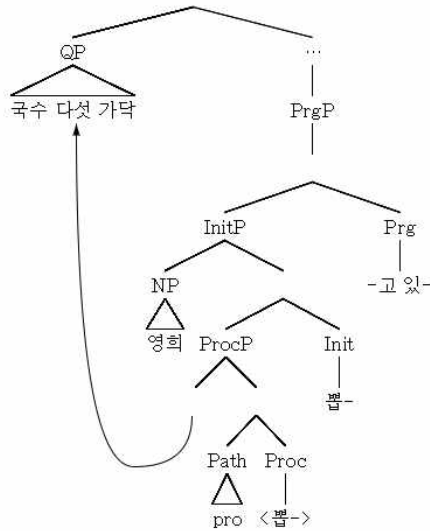
3.2.1에서 한 관찰에 근거하여 (19')를 설명하기 위한 다음의 가정을 제안한다.

(21)

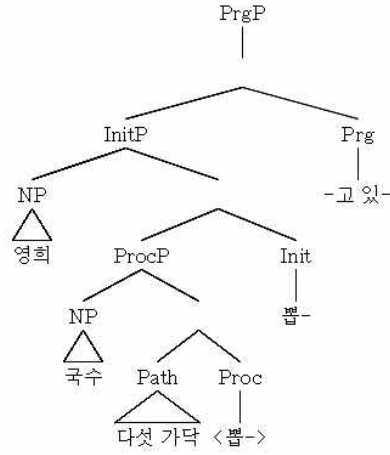
- 가. 진행상 표지 “-고 있-”은 진행사구(Progressive Phrase, 이하 PrgP)를 투사하고 InitP를 보충어로 선택한다.
- 나. 양화사로 타입 전환된 분류사구 및 측정 수량사는 양화사구(Quantifier Phrase, 이하 QP)를 투사한 채로 기저부에 삽입된다.
- 다. 양화사 섬 효과(island effect) $\ast QP_i$ [ProgP ... [Path ... $\&$...] ...]

(21)에 따라 (18'가), (18'나)는 LF에서 각각 다음의 수형도 (22가), (22나)를 가진다. (편의상 주어의 [Spec TP] 이동은 생략하였다)

(22) 가.



나.



(22가)에서 보듯이 QP로 타입 전환된 분류사구 “국수 다섯 가닥”은 기저에서 경로보다 높은 위치인 [Spec ProcP]의 위치에 (21다)의 섬 효과를 받지 않고 LF에서 통사론적으로 상당히 높은 QP 안착 지점 (예컨대 [Spec CP])으로 이동하기 때문에 적절한 작용역을 형성할 수 있다. 따라서 (19'가)와 같이 수사의 크기를 감소시키면 논리식에 존재하는 존재양화사 \exists 에 의하여 함의 관계를 유도한다. 그러나 (22나)에서 보듯 (18'나)의 측정 수량사는 QP로 타입 전환이 불가능하기 때문에 이동 없이 생성되었던 위치 그 자리에서 해석되어야 한다. 만약 QP로 삽입되었다면 경로보다 낮은 위치에 있기 때문에 QP 이동이 섬 효과 제약 (21다)를 어기게 되어 비문이 된다. 제자리에서 해석된 측정 수량사는 경로를 수식하는 역할을 하고 수사 감소에 의한 함의 관계가 유도되지 않는다.

참고문헌

- 김윤신(2006), “한국어 동사의 사건구조와 어휘상(語彙相)”, 『한국어학』 30, 31-61.
 김윤신(2007), “사건 함수로서의 상 보조 용언에 대한 연구”, 『한국어학』 35, 17-47.
 박소영(2009), “수량사 구성의 형태-통사론적 분석”, 『형태론』 11-1, 1-27.
 박소영(2017), “한국어 명사구 생략의 통사론 -수분류사 구성을 중심으로-”, 『언어와 정보 사회』 13, 391-422.
 서셋별(2011), 『Constraints on NP-ellipsis and DP-internal Movement』, 서울대학교 석사학위논문.
 채완(1982), “국어수량사구의 통시적 고찰”, 『진단학보』 53·54, 155-170.
 채완(1983), “국어 수사 및 수량사구의 유형적 고찰”, 『어학연구』 19-1, 19-34.
 최기용(2001), “한국어 수량사 구성의 구조와 의미: 비속격형을 중심으로”, 『어학연구』 37-3, 445-482.
 최기용(2009), 『한국어 격과 조사의 생성통사론』, 한국문화사.
 최형용 외(2017), 『한국어 분류사 연구』, 도서출판 역락.

- 홍사만(2008), “한·일어 분류사의 대조 연구”, 『언어과학연구』 44, 1-27.
- Aikhenvald, A. (2001), *Classifiers: A Typology of Noun Categorization Devices*, Oxford: Oxford UP.
- Chierchia (1998), Plurality of mass nouns and the notion of “semantic parameter”, In S. Rothstein (Ed.), *Event and Grammar* (pp. 53 - 104), Kluwer Academic Publishers.
- Landman, F. (2004), *Indefinites and Type of Sets*, Oxford: Wiley-Blackwell.
- Pustejovsky, J. (1991), The syntax of event structure, *Cognition*, 41, 47-81.
- Ramchand, G. (2008), *Verb Meaning and the Lexicon: A First Phase Syntax*, Cambridge: Cambridge University Press.
- Rothstein, S. (2011), Counting, Measuring and the Semantics of Classifiers, *The Baltic International Yearbook of Cognition, Logic and Communication*, 6, 1-42.
- Rothstein, S. (2016), Counting and Measuring, *The Baltic International Yearbook of Cognition, Logic and Communication*, 11, 1-49.
- Rothstein, S. (2017), *Semantics for Counting and Measuring*, Cambridge: Cambridge University Press.
- Saeed, S. (2016), Motion Events: Syntactic-Semantic Mapping, *American International Journal of Contemporary Research*, 6(3), 55-82.
- Saito, M., J. Lin, & K. Murasugi. (2008), N'-Ellipsis and the Structure of Noun Phrases in Chinese and Japanese, *Journal of East Asian Linguistics*, 17(3), 247-271.
- Sim, C. (2006), Measure Phrase and the -ka/-lul alternation in Korean, *University of Pennsylvania Working Papers in Linguistics*, 12(1), Article 26.

술어와 논항의 구체성 정도와 문자 그대로의 의미 판별 기준

임서현(서울대학교 언어학과, 자동화시스템연구소)

1. 서론

이 논문은 단어 의미의 구체성 정도(concreteness ratings)가 합성 표현의 문자 그대로의 의미(literal meaning) 판별의 기준이 될 수 있는지 영어 동사와 논항 결합 예를 통해 확인하고자 한다. 예를 들어 아래 제시한 (1)에서 (1a)는 문자 그대로의 의미이고, (1b)는 은유적 의미 확장으로 맥락적 의미라고 볼 수 있다.

- (1) a. 강물이 흐른다.
- b. 세월이 흘렀다.

위의 예에서 동사 ‘흐르다’가 ‘강물’이라는 명사와 쓰일 때 (1a)는 문자 그대로의 의미이지만, ‘세월’이 논항으로 올 때 (1b)는 시간의 진행을 물의 흐름에 빗대어 은유적으로 확장한 것이다.

흥미로운 것은 ‘강물’은 눈을 통해 시각적으로 지각이 가능한 구체적 대상인 반면, ‘세월’은 지각 가능한 구체적 대상이라기보다는 추상적 개념을 나타내는 단어라는 점이다. ‘흐르다’라는 구체적인 현상을 나타내는 술어와 구체적 대상을 지시하는 논항이 결합할 때는 문자 그대로의 합성적 의미가 도출되지만, 구체 술어 ‘흐르다’와 추상적 단어 ‘세월’이 결합할 때는 은유적 의미로 해석된다. 의미 타입으로 보면 ‘흐르다’는 ‘물’이나 ‘커피’, ‘피’ 등과 같이 흐를 수 있는 ‘액체’ 타입의 논항을 취한다는 타입 제약이 있다. ‘세월’이 논항으로 허가되는 것은 ‘세월’을 ‘액체’ 타입의 명사에 대응하도록 은유적으로 해석하기 때문이다. 이렇게 술어와 논항의 구체성 정도 매치를 확인하여 문자 그대로의 의미와 맥락적 의미를 구분할 수 있는 것으로 보인다. 문제는 모든 술어-논항 결합이 그러한 경향을 보이는가, 그리고 의미 타입(semantic type)보다 더 좋은 판단 기준이 될 수 있는가를 확인할 필요가 있다.

예초 구체성 정도를 통해 문자 그대로의 의미와 맥락적 의미 구분을 시도한 연구는 Turney et al. (2011)이다. 이들은 형용사-명사 쌍을 구체성 정도를 이용하여 문자 그대로의 의미와 맥락적 의미로 자동 분류하는 감독 모델(supervised model)을 개발하여 79% 정답률을 보이는 연구를 발표하였다. Hill & Korhonen (2014)는 구체성 정도와 주관성(subjectivity)이 형용사-명사 결합 의미에 영향을 미친다는 주장을 하였다. 이 논문은 형용사-명사 결합이 아니라 영어 동사를 중심으로 술어(동사)와 논항의 결합에 대해서도 구체성 정도가 문자 그대로의 의미와 맥락적 의미로 구분하는 기준이 될 수 있는지 논의한다.

먼저 2절에서는 이론적 배경으로 문자 그대로의 의미와 맥락적 의미 구분에 관한 의미, 화용론적 연구에서 제시된 주장들을 요약한 다음 단어 의미의 구체성/추상성에 관한 정의와 관련 논의를 간략히 소개한다. 이 논문에서는 문자 그대로의 의미를 전통적 의미론과 화용론의 견해에 따라 우리가 언어적 의미 혹은 단어의 어휘의미라 부르는 관습적 의미(customary sense)를 문자 그대로의 의미로 정의한다. 구체성에 관하여는 지각할 수 있는 대상을 지시하는 단어를 구체적인 단어로 언어적 설명으로 이해할 수 있는 단어를 추상적 단어로 인식한다. 구체적인 단어와 추상적인 단어는 이분법적으로 가릴 수 있는 것

이 아니라 구체성은 정도 차이를 갖는 스펙트럼으로 표현할 수 있고 각 단어는 그 스펙트럼 상의 어느 위치를 차지한다고 본다. 더 자세한 내용은 2절에서 설명한다. 3절에서는 이러한 이론적 배경을 토대로 영어의 ‘hope’, ‘put’, ‘meet’, ‘fall’, ‘smoke’의 구체성 정도와 논항의 구체성 정도 차이를 분석함으로써 문자 그대로의 의미와 맥락적 의미 구분에 기준으로 작용하는지 살펴보도록 한다. 마지막으로 이 논문의 결론으로 구체성 정도와 맥락적 의미 확장의 관계를 요약하고 앞으로 진행할 향후 연구 방향을 소개하면서 이 논문을 마치고자 한다.

2. 이론적 배경

전통적으로 단어의 어휘의미가 실제 사용에서 맥락에 의존하여 변하는 변이 유형은 어휘적 축소(lexical narrowing; e.g. *drink* ‘drink alcohol’), 근사적 사용(approximation; e.g. *square* ‘squarish’)과 은유적 확장(metaphorical extension; *nightmare* ‘bad dream’)으로 구분됐다. 그런데 관련성 이론(Relevance Theory)이 이러한 구분에 반대하면서 통합적 설명을 시도함으로써 문자 그대로의 의미와 비유적 의미(figurative meaning) 구분에 대한 논쟁이 시작되었다. 문자 그대로의 의미와 비유적 의미 구분은 의미론과 화용론, 진리 조건적 의미와 비진리 조건적 의미, 명시적(explicit) 의미와 암묵적(implicit) 의미 구분의 논쟁과 매우 밀접한 관련을 갖는다(Jaszczolt, 2016).

언어표현에 언어적으로 인코딩된 의미를 최소화하는 반면 맥락의 역할을 최대화하고, 의미론과 화용론의 구분에 반대하는 맥락주의(contextualism) 입장에서는 문자 그대로의 의미와 비유적 의미를 구분하고 전자로부터 후자로 변이하였다는 주장이 적절치 않다. 왜냐하면, 어휘의미는 미명세되어 있고, 맥락이 의미를 결정하기 때문이다. 그러므로 단어의 의미는 어떤 한 의미로부터 맥락에 의존하여 의미 변이를 한다기보다는 모든 언어표현의 의미는 동등한 자격으로 미명세되어 있고 동일하게 맥락이 주는 정보를 이용하여 명세된다. 문자 그대로의 의미를 인정한다고 하더라도 그 의미는 화자(speaker)의 의미를 나타낸다. 즉, 문자 그대로의 의미는 화자의 생각을 외적으로 표현한 것이다. 반면, 최소주의(minimalism)는 문자 그대로의 의미와 맥락적 의미 변이 즉, 비유적 의미 구분을 인정하면서 문자 그대로의 의미를 관습적 의미(customary sense), 다시 말해 우리가 언어적 의미라 부르는 것으로 여긴다.

문자 그대로의 의미에 관한 논쟁은 그 자체로 매우 흥미롭고 중요하지만, 이 논문의 핵심 주제는 아니므로 자세한 논의는 하지 않기로 한다. 다만, 구체성 정도와 관련하여 문자 그대로의 의미와 비유적 의미의 구분을 하고자 하는 이 논문의 의도를 실현하기 위해서는 위의 두 가지 입장 중 하나를 선택해야 하는데 이 논문에서는 전통적인 의미론의 입장을 따라 최소주의 입장 즉, 문자 그대로의 의미와 비유적 의미의 구분과 문자 그대로의 의미에서 비유적 의미로의 맥락의존적 의미 변이라는 입장을 택하기로 한다. 다음은 단어 의미의 구체성/추상성에 관한 논의를 간략히 소개한다.

임서현(2018)에서 기술한 바대로, 최근 들어 추상어휘의 의미 표상, 처리 등에 관한 연구의 필요성이 대두되면서 이론 언어학, 심리언어학, 신경과학뿐만 아니라 자연어처리, 인공지능, 로보틱스 등 다양한 분야에서 추상어휘에 관한 연구를 진행하고 있다. 애초에 언어표현의 구체성/추상성에 관한 연구의 동기는 심리언어학과 인지과학에서 제기되었다. 심리언어학적 동기는 이른바 “구체성 효과(Concreteness Effect)”라는 것으로 심리언어학에서 진행되는 언어 관련 대부분의 실험에서 구체적이라 여겨지는 언어표현이 더 신속하고, 정확하며, 좋은 결과를 낸다는 경험적 인식이다. 또 다른 중요한 동기는 체화 인지 이론(Embodied Cognition Theory)이 출현과 발전이다. 체화 인지 이론은 인간의 언어 습득과 처리, 이해 등이 언어 그 자체만으로 진행되는 것이 아니라 그 언어표현이 지시하는 행동이나 대상과 관련된 실세계와 인간의 상호작용이 영향을 미친다는 이론이다. 이를 기반으로 자연어처리와 인공지능, 로보틱스 등에서 다양성 모델과 발달 로보틱스가 출현하였는데, 이 또한 언어표현 의미의 구체성/추상성에 관한 연

구의 동력이 되었다.

구체성(혹은 추상성)을 어떻게 정의하는가?¹⁾ 어떤 단어가 추상어휘인지 구체어휘인지 결정하는 기준은 무엇인가? 추상과 구체를 구분하는 가장 분명한 특성으로 지각 가능성(perceivability)을 들 수 있다. 구체어휘는 감각을 통해 지각할 수 있는 대상을 지시하는 단어, 추상어휘는 물리적 경험으로 지각할 수 없는 대상들을 언어적으로 설명하는 단어로 정의된다(Paivio et al. 1968, Barsalou 2003). 그러나 사실 구체어휘와 추상어휘는 정확히 둘로 나눌 수 있는 것이 아니다. 오히려 단어의 구체성은 가장 추상적인 단어와 가장 구체적인 단어를 양극으로 하여 정도의 차이를 갖는 연속적 스펙트럼으로 표현할 수 있다. 예를 들어 사회적 역할을 지시하는 *physician*과 같은 단어는 단순한 객체를 가리키는 *bottle*보다 추상적이지만, *odd number*와 같이 순수하게 어떤 것을 정의하는 단어들보다는 덜 추상적이라 할 수 있다. 구체성은 단어가 감각을 통해 지각될 수 있는 정도를 말한다(Paivio et al. 1968).

이 논문에서는 구체성이 이분법이 아니라 정도의 차이를 나타내는 스펙트럼으로 표현된다는 주장을 받아들여 Brysbaert et al. (2014)가 순위매김 해 놓은 영어 단어 구체성 정도 데이터셋을 이용하여 문자 그대로의 의미(literalness)를 판별할 수 있는지 검토한다. 이 영어 단어 구체성 순위매김 프로젝트는 대략 40,000개의 미국 영어 단어와 두 단어 표현(예: *zoom in*)에 대해 클라우드 소싱을 통해 모국어 화자들에게서 구체성 정도 점수를 얻은 데이터셋이다. 이 연구는 참여자들에게 가장 추상적인 단어를 1, 가장 구체적인 단어를 5로 하는 5점 규모(scale)를 제시하고 그중 하나를 선택하도록 하여 구체성 정도를 점수로 제시한다. 구체어휘는 5가지 감각 경험을 통해 그 지시대상을 지각할 수 있는 단어로 정의되고, 추상어휘는 감각 경험을 통해 지각될 수 없고, 대신 다른 단어들에 의해 설명될 수 있는 단어들로 정의된다. 예를 들어 영어 동사 *climb*은 4.11로 매우 구체적인 어휘에 속한다. *believe*는 1.55로 매우 추상적인 의미를 지닌다고 보아야 한다(임서현, 2018).

구체성 정도와 문자 그대로의 의미가 관련성을 가지는 것은 서론에서 언급한 대로 Hill & Korhonen (2014)에서 그 근거를 찾을 수 있다. 그들은 형용사와 명사가 합성된 결합구조가 문자 그대로의 의미를 갖는가 아닌가를 판단하는데 형용사의 구체성/추상성과 명사의 구체성/추상성이 영향을 미친다고 주장한다. 즉, 추상적인 형용사가 추상적인 명사와 결합하거나 구체적인 형용사가 구체적인 명사와 결합을 하는 경우는 문자 그대로의 의미를 전달하지만, 형용사와 명사의 구체성과 추상성이 엇갈릴 때는 문자 그대로의 의미가 아닌 맥락적 의미 변이를 한 것으로 판단한다. 예를 들어 구체적인 색깔을 나타내는 형용사 ‘white’는 구체성이 강한 단어인데 이 단어가 ‘table’과 같은 구체적인 명사와 결합할 때 ‘white table’은 문자 그대로의 의미로 하얀색 테이블을 지시하지만, ‘hope’처럼 추상적인 단어와 결합할 때는 “하얀색 희망”이 아니라 “유망주”를 의미한다. 이때에는 문자 그대로의 의미라 할 수 없다. 이러한 관찰을 문자 그대로의 의미 사용 추출 기준으로 적용할 수 있는지 살펴볼 것이다. 다음 절에서는 영어의 동사-논항 결합을 이 기준으로 분석해 보도록 한다.

3. 구체성 정도와 문자 그대로의 의미

Brysbaert et al. (2014)의 구체성 정도 주석 데이터셋은 영어 명사, 동사, 형용사에 대해 구체성 정도를 1-5 규모 스펙트럼 중 점수를 매기도록 하였다. 이 주석은 단어의 다의성(polysemy)을 고려하지 않는 대신 동사의 원형과 과거형, 완료형, 진행형 4가지 형태 각각에 대해 따로 구체성 정도를 주석한 것이다. 이 논문에서 구체성 정도를 이용한 문자 그대로의 의미 사용과 비유적 의미 사용 구분을 위하여 영어 동사 중 가장 추상적인 동사부터 가장 구체적인 동사까지 각 레벨에 속하는 동사 5개를 선택하였다. 아래 (2)에 그 동사 목록과 구체성 정도를 제시한다.

1) 추상성과 구체성은 상보적인 개념이므로 이 논문에서는 ‘구체성’으로 통일하여 사용한다.

- (2) a. hope (1.25), hoped (1.69)
- b. put (2.5)
- c. meet (3), met (2.58), meeting (3.31)
- d. fall (4.04), fell (3.55), fallen (3.53), falling (3.64)
- e. smoke (4.96), smoked (3.65), smoking (4.32)

위 (2)에서 ‘hope’가 가장 추상적인 동사로 구체성 정도 1.25이고, 가장 구체적인 동사는 ‘smoke’로 5점 만점에 4.96이다. ‘put’, ‘meet’, ‘fall’은 중간에 있는 동사로 직관적으로 보아도 동사의 의미가 지각 가능한 지시대상 행위가 쉽게 떠오르는 구체성이 있으면서도 추상적인 의미로도 쓰이는 것을 알 수 있는 단어이다. 특히 이 동사들은 워드넷에서 ‘put’은 9개, ‘meet’은 13개, ‘fall’은 32개로 의미를 나누어 놓은 것을 보면 알 수 있듯이 다의성이 강한 동사들이다. 이제부터 각 동사의 워드넷 신셋별 예문을 이용하여 그 예문에서 논항들의 구체성 정도와 동사의 구체성 정도의 결합에 대해 살펴보자. 먼저 가장 구체성 정도가 높은 ‘smoke’부터 아래 예문과 구체성 정도, 그리고 논항과 동사의 구체성 정도 차이를 제시한다.

- (3) 동사 ‘smoke’의 워드넷 신셋별 예문과 구체성 정도
 - a. We never **smoked** marijuana. [marijuana (4.89)]
 - b. The chimney was **smoking**. [schimney (5)]

동사 ‘smoke’는 4.96으로 가장 구체적인 동사 중 하나이다. ‘smoked’는 3.65, ‘smoking’은 4.32이다. 워드넷은 ‘smoke’의 신셋을 (3)에서 제시한 대로 2개로 나누어 제시한다. 두 가지 예 모두 논항도 구체적인 대상을 지시한다. 따라서 동사와 논항 모두 구체성 정도가 높아 ‘smoke’는 은유나 환유 등에 의한 의미 확장이 진행되지 않고 문자 그대로의 의미로 사용된 것으로 보인다. 보충하자면, ‘smoke’의 디폴트 논항 선택제약은 추상적인 논항을 허가하지 않는데 그것이 보편적인 의미 확장 방식인 은유나 환유 등에 의한 확장이 사전에 등록될 정도로 자주 쓰이지는 않은 것으로 보인다. 다음으로 위에 제시한 5개의 동사 중 가장 구체성이 약한 ‘hope’의 예를 보자.

- (4) 동사 ‘hope’의 워드넷 신셋별 예문과 구체성 정도
 - a. I hope she understands that she cannot expect a raise. [I (3.93)]
 - b. I hope that all will turn out well. [I (3.93)]
 - c. I hope to have finished this work by tomorrow. [I (3.93)]

영어 동사 ‘hope’는 1-5 규모 구체성 정도에서 가장 추상적인 동사 범주에 속하는 1점대로 그중에도 구체성 정도 1.25의 매우 추상적인 동사이다. ‘hope’는 특히 심리동사이므로 전형적인 주어 논항은 심리적 경험이 가능한 ‘인간’으로 경험주 의미역을 갖는다. 주어 논항의 구체성 정도는 높은 편이지만, ‘hope’가 워낙 추상적이고 목적어 논항으로 주로 선택되는 명제나 사건 지시 논항은 추상적이므로 전체 문장의 구체성 정도는 낮다고 보아야 한다. 위에 제시한 예들은 주어 논항 ‘I’의 구체성 정도가 3.93으로 높은 편으로 추상적인 단어라기보다는 구체적 단어로 보아야 함에도, 술어와 논항의 구체성 정도의 불일치에 따른 비유적 의미 사용의 예로 보이지는 않는다. 그러나 다음 예를 보면 상황이 조금 달라진다.

- (5) a. The woman **hopes** to eat the banana. [woman (4.46)]
- b. The rat **hopes** to escape from the maze [rat (4.85)]

c. The newspaper **hopes** to have eight times as many articles with comments by the year's end.

[newspaper (4.82)]

위의 예를 보면, 'woman'은 구체성 정도가 4.46이고, 'hope'는 1.25로 그 정도 차이가 3.21이나 되지만, (5a) 문장을 비유적 의미로 보기는 어렵다. 오히려 'woman'은 'hope'의 전형적인 주어 논항이라고 볼 수 있다. 따라서 술어와 논항의 구체성 정도 차이가 크지만, 문자 그대로의 의미로 해석되므로 Turney et al. (2011)이나 Hill & Korhonen (2014)에서 주장한 구체성 정도와 문자 그대로의 의미 간의 관계를 (5a)에 적용하기는 어렵다. 오히려 이 예는 구체성 정도 불일치보다 술어와 논항의 의미 타입 매치가 문자 그대로의 의미 판별에 더 타당한 기준임을 보여준다.

반면, 'rat'과 'newspaper'는 둘 다 4.8을 넘는 매우 구체적인 단어로 추상적 단어 'hope'와 구체성 정도 불일치가 있고, (5b)와 (5c)는 비유적 의미로 해석된다. 'rat'은 일반적으로 희망할 수 있는 주체가 아니라 이 예문의 의미는 'rat'을 의인화한 것으로 보는 것이 더 적절할 것이다. 보통 'newspaper'는 인공물 (artifact)로 희망할 수 있는 주체가 아니라 사물이다. 그러나, (5c)는 비문이 아니며(이 예문은 구글 검색을 통해 얻은 예이다), 'newspaper'는 신문이 아니라 신문사 혹은 신문 기자들로 해석되는 것이 적절하다. 이 경우는 술어는 매우 추상적인데 논항은 매우 구체적이므로 구체성 정도가 문자 그대로의 의미 판별 기준으로 적절해 보인다. 그러므로 이 예들은 술어와 논항의 구체성 정도 차이가 맥락적 의미 변이의 기준이 된다는 것을 입증하는 예이다.

이때에도 당연히 의미 타입 불일치는 발생한다. 다시 말해 이 예들에서 논항의 의미 타입은 술어가 요구하는 의미 타입 제약에 어긋난다. 역으로 술어도 논항 명사의 의미 타입 선택제약에 어긋난다. 한 가지 흥미로운 것은 (5a)의 논항이 전형적인 'hope'의 주어 논항이라 할 때 (5b)와 (5c)의 술어-논항 구체성 정도 차이는 (5a)의 정도 차보다 크다는 점이다. 어쩌면 구체성 정도 차이가 문자 그대로의 의미 판별 기준이 되기 위해서는 전형적인 술어-논항 의미 타입(혹은 구체성 정도 차이)보다 큰 폭이어야 한다고 주장할 수도 있을 것이다. 이를 증명하기 위해서는 전형적인 술어-논항 의미 결합 제약이 결정되어야 하고, 그 구체성 정도 차이가 결정되어야 한다.

지금까지 양극단의 술어 2개를 살펴보았다. 이제 중간 정도의 구체성을 갖는 'meet', 'put', 'fall'에 대해 보도록 하자. 먼저 'put'의 예를 보자.

(6) 동사 'put'의 워드넷 신셋별 예문과 구체성 정도

- a. **Put** your things here. [thing (3.17), here(3.13)]
- b. That song **put** me in awful good humor. [song (4.46), me (4.33), humor (2.26)]
- c. She **put** too much emphasis on her the last statement. [she (3.36), emphasis (2.12), statement (2.96)]
- d. **Put** money into bonds. [money (4.54), bonds (3.62)]
- e. We **put** the time of arrival at 8 P.M. [time (3.07)]
- f. He **put** her to the torture. [her (3), torture (3.59)]
- g. **Put** these words to music. [word (3.56), music (4.31)]
- h. I **put** these memories with those of bygone times. [I (3.93), memory (2.38)]

영어 동사 'put'은 대표적인 이동동사 중 하나로 동사가 지시하는 이동 사건을 지각할 수 있다는 의미에서 구체적인 의미가 있는 동사로 분류될 수 있지만, 대부분의 이동동사가 그렇듯이 추상적인 논항과 함께 쓰여 추상적 의미로 맥락적 의미 확장이 되면서 다의성이 큰 단어이다. Brysbaert et al. (2014)가 분석한 'put'의 구체성 정도는 2.5로 정확히 중간 정도의 구체성을 갖고 있다. 워드넷은 (6)에서 보듯이 9개

의 신셋을 제시한다. 문제는 워드넷의 신셋 구분이 구체성의 정도를 구분해 주지 못한다는 것이다. 예를 들어 (6a)의 경우 그 정의가 “put into a certain place or abstract location”으로 추상적 장소와 구체적 장소를 구분하지 않고 넣는 행위에 초점을 맞춘 의미이다. 그러므로 워드넷 신셋별로 구체적 의미인지 추상적 의미인지도 구분이 되지 않을뿐더러 문자 그대로의 의미인지 비유적 의미인지도 구분하기 쉽지 않다. 이는 사실 워드넷의 너무 세밀하며 명확하지 못한 의미 구분 기준이 원인이라 보아야 할 것이다. 워드넷에 대한 이러한 비판은 이미 잘 알려진 사실이다. 이를 염두에 두고 동사 ‘put’과 목적 논항의 구체성 정도 차이를 비교해 보면 다음과 같다.

h(0.12) put these memories < c(0.38) put emphasis < f(0.5) put her < e(0.57) put time < **a(0.67) put things** < g(1.06) put words < b(1.83) put me < d(2.04) put money

가장 작은 정도 차이를 보이는 예는 (6h) “put these memories”이고, 가장 큰 차이를 보이는 예는 (6d) “put money”이다. 그러나 이 예들은 모두 문자 그대로의 의미라기보다는 비유적 의미로 은유적 의미 확장이라 볼 수 있다. 결국, 이 예들이 구체성 정도의 차이가 문자 그대로의 의미 사용의 판별 기준으로 타당치 않음을 보여주는 것이다. 직관적으로 판단할 때 가장 문자 그대로의 의미로 이동동사의 예로 쓰인 예는 (6a)의 “put things”이다. 이제 워드넷의 예들보다는 더 명확한 문자 그대로의 의미와 맥락적 의미 확장의 대립을 보여주는 예들을 보자.

- | | |
|----------------------------------------------------------------------------------|------------------------------------------|
| (7) a. He put his <u>hand</u> in his <u>pocket</u> . | [put (2.5), hand (4.72), pocket (4.68)] |
| b. I put <u>them</u> on the <u>chairs</u> . | [put (2.5), them (3.04), chair (4.58)] |
| c. The Senate had never put any <u>limits</u> on <u>nuclear testing</u> . | [put (2.5), limit (2.28), testing(3.43)] |
| d. That will put a <u>brake</u> on <u>industrial well-being</u> . | [put(2.5), brake (4.44), being(1.93)] |

위의 예들을 보면 직관적으로 (7a, b)는 문자 그대로의 의미로 이동의 의미로 쓰였지만, (7c, d)는 장소 이동을 추상적 장소로 은유적 확장을 한 예이다. ‘nuclear testing’과 ‘well-being’은 추상적 장소의 의미로 비유적 의미 사용을 한 것이고, ‘limit’과 ‘brake’도 이동하는 대상에 비유하여 쓰인 것이다. 각 목적어(대상) 논항과 장소 논항의 구체성 정도는 (7a, b)와 (7c, d)를 비교하면 꽤 큰 차이가 있음을 알 수 있다. 결론적으로 동사의 구체성 정도가 중간 정도일 때 그 의미의 다의성이 커지며²⁾, 그때는 논항의 구체성 정도가 술어-논항 결합의 문자 그대로의 의미 판단 기준이 될 수 있다³⁾.

다음으로 구체성 정도가 3인 ‘meet’을 보자. 워드넷은 이 동사의 신셋을 13개로 제시하고 있다. 다음 예들은 신셋별로 제시된 것으로 예문이 없는 경우는 제외하였다.

- (8) 동사 ‘meet’의 워드넷 신셋별 예문과 구체성 정도
- How nice to **meet** you again. [you: 4.11]
 - The lines **meet** at this point. [line: 4.86]
 - Does this paper **meet** the requirements for the degree? [requirement: 2.52]
 - meet** a need [need: 1.69]

2) 물론 역으로 다의성때문에 언어 화자들이 구체성 정도를 중간 정도로 파악한다고 설명할 수도 있겠다. 즉, 다의성이 큰 단어들에 대해서 문자 그대로의 의미로 쓰인 예와 비유적 의미로 쓰인 예들을 언어 화자들이 직관적으로 파악하고 기억하고 있으므로 구체성 정도가 중간 정도로 추측되는 것으로 이해할 수도 있다.

3) 어떤 논항의 구체성 정도가 전체 합성 표현의 문자 그대로의 의미 판단에 결정적인 영향을 미치는지는 더 연구해 보아야 한다.

- e. I **met** this really handsome guy at a bar last night. [guy: 4.68]
- f. Let's **meet** in the dining room [us: 3.59]
- g. Can you **meet** me at the train station? [me: 4.33]
- h. My proposal **met** with much opposition. [proposal: 2.75, opposition:2.25]
- i. **meet** a violent death [death: 4.22]
- j. The two buildings **meet**. [building: 4.64]

동사 ‘meet’이 가장 많이 쓰이는 전형적인 예는 (8a, e, f, g)처럼 사람들끼리 만나는 경우이다. 이런 의미는 문자 그대로의 의미로 사용된 것이다. 그 외에 위의 예들에서 논항의 구체성 정도가 높아서 술어와 논항의 정도 차가 큰 (8b, i, j)와 같은 예들은 만나는 대상으로 사물을 의인화한 것으로 의미 변이를 한 것이다. (8c, d, h)처럼 구체성 정도가 낮은 추상적 단어를 논항으로 취하는 경우는 은유적 의미 확장으로 이때도 문자 그대로의 의미 사용은 아니다. 이 예들도 술어와 논항의 구체성 정도 차가 클수록 비유적 의미로서의 쓰임이 강하다는 주장에 반하는 예들이다. 따라서 술어와 논항의 구체성 정도 차이를 문자 그대로의 의미 판단 기준으로 쓰기는 모호하다.

마지막으로 구체성 정도가 4.04인 ‘fall’은 32개의 워드넷 신셋을 갖고 있지만, 이동동사의 하나로 분류되며 물리적 하강 이동을 나타내는 구체 동사로 보는 것이 직관적으로 가장 일반적인 의미이다. 눈여겨 볼 만한 것은 앞서 제시한 ‘put’도 이동동사이지만 구체성 정도가 2.5로 낮은 편인 데 반해 ‘fall’은 훨씬 더 구체성 정도가 크다는 점이다. 아래 워드넷 신셋 예문들을 제시한다.

- (9) 동사 ‘fall’의 워드넷 신셋 예문과 구체성 정도
- a. The branch **fell** from the tree. [branch: 4.9, tree: 5]
 - b. The unfortunate hiker **fell** into a crevasse. [hiker: 4.53, crevasse:]
 - c. The barometer is **falling**. [barometer: 4.57]
 - d. **Fall** in love [love: 2.07]
 - e. **fall** asleep [asleep: 3.71]
 - f. **fall** into a trap [trap: 4.3]
 - g. She **fell** ill. [she: 3.36, ill: 3.21]
 - h. They **fell** out of favor. [they: 2.93, favor: 2.07]
 - i. **fall** prey to an imposter [prey: 3.61, imposter: 3.38]
 - j. **fall** into a strange way of thinking [way: 2.34]
 - k. She **fell** to pieces after she lost her work. [She: 3.36, piece: 4.14]
 - l. **fall** into a category [category: 2.82]
 - m. Rain, snow, and slee were **falling**. [rain: 4.97, snow: 4.85, slee:]
 - n. we must stand or **fall** [we: 3.08]
 - o. **fall** by the wayside [wayside: 2.75]
 - p. The cabin pressure **fell** dramatically. [pressure: 3.07]
 - q. Her weight **fell** to under a hundred pounds [weight: 3.94, pound: 4.61]
 - r. His voice **fell** to a whisper [voice: 4.13, whisper: 4.34]
 - s. Many soldiers **fell** at Verdum [soldier: 4.72, Verdum:]
 - t. Several deer have **fallen** to the same gun. [deer: 4.86, gun: 4.83]
 - u. The shooting victim **fell** dead. [victim: 3.59, dead: 4.2]
 - v. Light **fell** on her face. [light: 4.21, face: 4.87]
 - w. The cities **fell** to the enemy [city: 4.79, enemy: 2.83]
 - x. Christmas **falls** on a Monday this year. [Christmas: 3.41, Monday: 2.21]
 - y. The accent **falls** on the first syllable. [accent: 3.26, syllable: 3.26]
 - z. Adam and Eve **fell**. [Adam and Eve]
 - a'. The government **fell** overnight. [government: 2.88]
 - b'. The Qing Dynasty **fell** with Sun, Yat-sen. [dynasty: 2.46]
 - c'. The most difficult task **fell** on the youngest member of the team. [task: 2.84, member: 3.87]
 - d'. The onus **fell** on us. [onus: , us: 3.59]
 - e'. The pressure to succeed **fell** on the youngest student. [pressure: 3.07, student: 4.92]

위의 예에서 문자 그대로의 의미로 쓰인 예는 (9a, b, c, e, f, m)이다. 다른 예들을 보면 논항의 구체성 정도가 4를 넘게 높아도 문자 그대로의 의미로 쓰였다고 보기는 어려운 예들이 많다. 그러므로 구체성 정도가 ‘fall’의 문자 그대로의 의미와 비유적 의미(혹은 맥락적 의미) 확장의 구분을 위한 기준으로 작용한다고 보기는 어렵다. 우리는 이 절에서 5개 영어 동사와 그 논항의 합성 구성에 대해 워드넷 신셋별 예문을 이용하여 동사의 구체성 정도와 논항의 구체성 정도의 차이가 문자 그대로의 의미를 판별하는 기준이 될 수 있는지를 살펴보았다. 결론적으로 검토한 동사들 모두 문자 그대로의 의미와 맥락적 의미 변이를 가르는 기준으로 동사와 논항의 구체성 정도 차이가 일관되지 않음을 보였다. 이제 결론에서 그 원인에 대해 정리하고 향후 연구 방향을 제시하면서 이 논문을 끝내고자 한다.

4. 결론

우리는 지금까지 Brysbaert et al. (2014)의 구체성 정도 주석을 기반으로 위의 다섯 개의 대표적인 영어 동사와 논항의 구체성 정도 차이를 살펴보고 그 정도 차가 문자 그대로의 의미와 맥락적 의미 확장의 구분을 위한 기준으로 작용할 수 있는가에 대해 논의하였다. 결론은 구체성 정도의 차이는 문자 그대로의 의미 판별 기준으로 적절치 않다는 것이다. 사실 Turney et al. (2011)이나 Hill & Korhonen (2014)의 연구는 영어 형용사와 명사의 합성에 관한 것이었고, 이에 대해 추상적 형용사와 추상적 명사, 구체적 형용사와 구체적 명사의 결합이 문자 그대로의 의미를 유도하고, 이 둘의 결합이 어긋날 때 맥락적 의미 변이를 가져온다는 것이 그들의 연구 결과이다.

이 논문에서 시도한 것은 술어 특히 동사와 그 논항 - 주어, 목적어, 전치사 목적어 등을 포함하여 -의 구체성이 문자 그대로의 의미인가를 판별하는 데 있어 기준으로 작용할 수 있는가를 검증하는 것이었다. 영어 동사 5개 - 'hope', 'put', 'meet', 'fall', 'smoke' - 를 검토한 결과는 동사와 논항의 구체성 정도 차이가 그 기준으로 적절치 않다는 것이다. 그 이유는 아직 확실치 않지만, 몇 가지를 언급해 볼 수는 있겠다. 우선 형용사 - 명사 결합과 비교하면 동사 - 논항 결합은 훨씬 더 복잡하여 간단히 구체성 정도로 의미 확장을 설명할 수 없기 때문이다. 둘째로 생각해 볼 수 있는 것은 구체성 정도가 연속적이기 때문에 그에 따라 구체성 정도의 차이도 연속적이고, 따라서 그를 기반으로 하는 문자 그대로의 의미와 맥락적 의미 변이도 이분법적으로 구분될 수 있는 것이 아니라 정도 차를 갖는 연속적 스펙트럼으로 나타내야 한다는 점이다. 셋째, Brysbaert et al. (2014)의 구체성 정도 주석이 동사의 다의성을 고려하지 않은 채로 조사한 것이어서 구체적 뜻과 추상적 뜻이 혼합된 양상을 보인다는 것이다. 예를 들어 'put'이나 'meet'과 같은 동사도 구체적 의미로 쓰이는 것이 더 기본적인 의미로 생각되는데 구체성 정도 주석 조사에서는 오히려 추상적이거나 적어도 중간 정도의 의미가 있는 것으로 주석되었다는 것이 문제일 수 있다. 마지막으로 구체성 정도도 결국 술어와 논항의 결합에 관한 제약을 나타내는 것이라면 구체성보다는 의미 타입이 더 세밀하고 정확한 제약 조건을 드러내는 것으로 생각된다.

앞으로 연구를 통해 더 많은 술어-논항 결합 데이터를 검토해 봐야 구체성 정도 차이와 문자 그대로의 의미의 연관성을 더 정확히 파악할 수 있을 것이다. 구체성 정도 주석도 단어의 다의성을 고려한 주석이 요구된다. 또한, 구체성 정도와 함께 다른 요소들이 필요한지도 살펴봐야 한다. 술어와 논항의 의미 타입 매치가 강력한 판별 기준이 될 것은 확실하다. 그러나 술어의 제약과 논항의 제약이 충돌할 때 어떤 것이 먼저인가 하는 문제가 발생하게 된다. 결국, 분포주의적 관점을 피하고자 한다면 이 우선순위에 관한 연구도 필요할 것이다. 적절한 판별 기준이 구체화된다면 이를 바탕으로 자동 맥락적 의미 변이 판별 시스템 개발이 가능할 것이다. 이 논문의 가치는 술어와 논항의 구체성 정도 차이와 문자 그대로의 의미 사이의 관련성에 관한 연구를 처음으로 시도하였다는 것에 있다.

참고문헌

- 임서현(2018), "추상어휘 연구의 비판적 검토와 한국어 추상어휘 연구의 필요성", 『언어학』 82.
- Barsalou, L. (2003), Abstraction in Perceptual Symbol Systems. *Philosophical Transactions of the Royal Society B: Biological Sciences* 358(1435), 1177-1187.
- Brysbaert, M., Warriner, A., & V. Kuperman (2014), Concreteness Ratings for 40 thousand Generally Known English Word Lemmas, *Behavior Research Methods* 46(3), 904-911.

- Hill, F. & A. Korhonen (2014), Learning Abstract Concept Embeddings from Multi-Modal Data, Since You Probably Can't See What I Mean, in the *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP) 2014*.
- Paivio, A., J. Yuille, & S. Madigan (1968), Concreteness Imagery and Meaningfulness Values for 925 nouns, *Journal of Experimental Psychology* 76, 1-25.
- Turney, P., Y. Neuman, D. Assaf, & Y. Cohen (2011), Literal and Metaphorical Sense Identification through Concrete and Abstract Context, in the *Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing, EMNLP 2011*, 27-31 July 2011, John McIntyre Conference Centre, Edinburgh, UK, A meeting of SIGDAT, a Special Interest Group of the ACL: 680-690.
- Jaszczolt, Kasia M. (2016), *Meaning in Linguistic Interaction*, Oxford University Press.

<감사의 글>

이 논문은 2017년도 정부(과학기술정보통신부)의 재원으로 정보통신기술진흥센터의 지원을 받아 진행된 연구(2017-0-01780, 비디오 이해를 위한 이벤트-상황 지식체계 학습 및 이벤트 인식/관계추론 기술 개발) 임을 밝혀 둔다.

한국어 동사 의미의 구체성/추상성 정도에 대한 연구*

박신영(서울대학교 언어학과)

1. 연구 목적

본 연구는 기존에 논의되지 않았던 한국어 동사의 구체성에 대해 Brysbaert et al.(2014)의 영어 어휘 구체성 정도 연구를 바탕으로 논의해 보고자 한다. 이를 위해 구체성/추상성에 대한 개념을 정리하고, 영어 어휘의 구체성 정도 연구를 살펴보고, 이를 한국어 동사를 대상으로 적용하여 구체성 정도에 대해 비교 논의하고자 한다.

2. 연구 배경

- 구체어휘: 보통 ‘사과’, ‘창문’과 같이 상황 안에 직접적으로 감각을 통해 지각되거나 연결 지을 수 있는 상징물이 있다.

추상어휘: ‘평화’와 같이 언어를 통해서만 이해되어질 수 있다.

- 실제적으로 구체성과 추상성은 이분법적으로 명확하게 나뉘기보다는 지속적 스펙트럼 상에서 정도의 차이를 보임

- ‘먹다’ → ‘음식을 입에 넣고, 씹고, 삼키는’ 더 구체적인 활동들을 포함.

- 추상적인 개념 또한 환경과 지각이 상호작용하는 구체적인 상황들로 쪼개거나 재해석

예) ‘분노’의 경우 냄비에서 물이 끓어 넘치는 상황과 같은 경우로 유추(직접적인 경험에 의한 해석에 의존)

본 연구에서는 오감을 통해 경험되어 질 수 없거나 지시체가 없는 추상어휘 의미보다도 ‘먹다’와 같이 구체성을 띄는 동사들이 더 구체적인 사건들로 쪼개질 수 있는 점에 주목하며, 이러한 동사들의 구체성 정도에 대해 논해 보고자 한다.

* 이 논문의 연구내용은 “비디오 튜링 테스트(Video Turing Test; VTT): 인간 수준의 비디오 이해 지능 및 검증 기술 개발 과제”의 일부로 미래창조과학부의 지원을 받아 연구되었음을 밝힌다.

3. 연구 방법

영어 어휘 구체성 정도 연구(Brysbaert et al., 2014)에서는 영어 어휘의 구체성 정도를 성인 영어 모국어 화자를 대상으로 설문조사하여 데이터베이스를 구축하였다. 여기서 단어의 추상적이거나 구체적인 개념은 이분법적인 구분이 아닌 연속체 개념이며 이를 실제적인 수치로 보여 주었다. 이러한 연구를 바탕으로 한국어 동사 구체성 정도 연구와 이에 대한 데이터베이스 구축의 필요성을 논하고, 이를 위해 한국어에 적용하여 모의 실험한다.

- 연구대상: 빈도수와 영어 구체성 정도 점수별 분포를 고려하여 뽑은 동사 의미에 대응하는 110개의 한국어 동사

- 방법: 제시된 동사들에 대해 구체/추상 정도를 1(추상적)에서 5(구체적)까지의 점수 중 하나에 경험적이고 직관적인 선택에 따라 표시하도록 성인 한국어 모국어 화자에게 설문 조사한다.

실험 결과 분석 및 논의

• 구체성 정도 점수별 동사 분류

구체성 점수	동사
1	신뢰하다, 가정하다, 믿다, 원하다, 사랑하다 등
2	알다, 선호하다, 제한하다, 느끼다, 동정하다 등
3	머물다, 구하다, 찾다, 놀다, 꾸짖다 등
4	만들다, 퍼지다, 다치다, 사다, 보다 등
5	x

• 영어와 한국어 구체성 정도 비교 논의

- 한국어 동사가 영어보다 구체적인 경우

예) 만들다 4 - make 2.67

→ 영어의 경우 보다 다양한 의미로 사용되는 점이 반영된 것으로 보임.

예) make a cake, make money, make friends, etc.

- 영어 동사가 한국어보다 구체적인 경우 예) 따르다-pour

- 영어와 한국어 의미간 일대일 대응이 안되는 경우

예) plug(3.59) - 메우다(4.48)/막다(3.87), put(2.5)/lay(3.86) - 놓다(4.13)

남은 과제

- 한국어 동사 목록을 확대하고 영어 구체성 정도 실험 같은 방식의 실험 진행 필요.

- 동사가 갖고 있는 여러 의미에 따라 구체성의 정도는 달라질 수 있으므로 다의어 구분 필요.

참고 문헌

- 에반 톰슨(2016), 「생명 속의 마음」, 도서출판 b.
- 임서현(2018), 추상어휘 연구의 비판적 검토와 한국어 추상어휘 연구의 필요성, 「언어학」 82호(발간예정)
- G. 레이코프 & M. 존슨(2006), 「삶으로서의 은유」, 도서출판 박이정.
- Binder, Jeffrey R. et al. 2016. Toward a brain-based componential semantic representation, *Cognitive Neuropsychology*. 33:130-174
- Brybaert, Marc et al.(2014), Concreteness ratings for 40 thousand generally known English word lemmas, *Behavior Research Methods*. 46:904-911.
- Rasheed, N. et al.(2015), Acquisition of abstract words for cognitive robots. *Jurnal Teknologi*. 72(2):21-28.

2018 서울 언어학 포럼 발표 논문집

2018년 12월 6일 인 쇄

2018년 12월 8일 발 행

편 집 사단법인 한국언어학회
총무간사 김민하
발 행 사단법인 한국언어학회
회장 최성호

인 쇄

연락처: 사단법인 **한국언어학회**

서울특별시 관악구 관악로 1

서울대학교 인문대학 언어학과 사무실 내

우편번호: 08826 전화번호: 010-5319-1397

※ 이 책은 팔지 않습니다. ※