

Direct Perception, Cue Reliability and Caseless Subjects: Evidence from Conversational Korean*

Hanjung Lee
(Sungkyunkwan University)

Lee, Hanjung. 2021. Direct Perception, Cue Reliability and Caseless Subjects: Evidence from Conversational Korean. *Korean Journal of Linguistics*, 46-4, 987-1013. Case-marked subjects in Korean have been shown to differ in a systematic way from their caseless counterparts as to their interpretation. This paper examines a hitherto unexplained property of caseless-subject clauses triggering a direct perception interpretation. I first present evidence from analyses of conversation data demonstrating that caseless subjects most productively occur in clause types with a subject that is identifiable in the here and now as an agent directly by the speaker and/or the hearer. Based on this evidence, I propose a new account of Differential Subject Marking (DSM) in terms of cue reliability, arguing that the association of caseless subjects with direct perception in the here and now, agentivity and tense deficiency follows from an economical use of formal particles motivated by reliability of semantic and situational cues for identifying an argument function and tense. These preliminary results support efficiency-based accounts of case marking and grammar (Hawkins 2004, Haspelmath 2008, Jaeger 2010, H. Lee 2010, 2016, Lestrade & de Hoop 2016, Levshina 2021). (Sungkyunkwan University)

Key words: caseless subject, cue reliability, Differential Subject Marking, direct perception in the here and now, economy, tense deficiency

1. Introduction

A phenomenon that has attracted considerable interest in literature on case marking over the last decades is Differential Case Marking (DCM), in which some arguments are marked with case markers, but not others, depending

* Versions of this paper were presented at the 22nd Seoul International Conference on Generative Grammar (August 2020) and the Workshop on Data-oriented Approaches to Meaning in Korean and Japanese (October 2021). I am grateful to the participants in these two conferences and three anonymous reviewers for their useful comments. Any errors and misinterpretation are due to my own shortcomings.

on semantic and pragmatic features of the argument (Bossong 1985, Aissen 2003, de Hoop & de Swart 2008, Levshina 2021). An interesting and challenging problem for theoretical approaches to DCM is that caseless and case-marked arguments systematically contrast as to their interpretation. In this paper, I will focus on a hitherto unexplained difference between caseless and case-marked subjects in Korean first noted by Kwon & Zribi-Hertz (2008). Examples in (1) illustrate the relevant contrast among the two subject forms.

- (1) a. Pesu-ka o-n-ta.
 bus-Nom come-Pres-Decl
 (i) (What is going on?/Why am I hearing strange noises?)
 'There is the/a bus coming.'
 (ii) (Is the taxi coming?) 'No, the bus is coming.'
 b. Pesu o-n-ta.
 bus come-Pres-Decl
 (The speaker is looking at coming of the bus at the bus stop and
 talks to the hearer over the phone.)
 'Here comes the bus.'

In both the *-ka*-subject clause in (1a) and the caseless-subject clause in (1b), the coming of the bus is assumed to be new information for the addressee. But the two clauses differ with respect to the availability of a direct perception interpretation. As noted by Kwon & Zribi-Hertz (2008), the caseless-subject clause in (1b) is understood as describing a situation that the speaker is observing directly at the moment of utterance: the speaker utters (1b) standing at the bus stop, while watching the bus approaching. The nominative-subject clause in (1a) is on the other hand understood as talking about a situation that is not necessarily based on the speaker's direct perception: it is understood as a statement of the coming of the/a bus as a future event or as an ongoing event that is not necessarily observed directly by the speaker or the hearer. This contrast in interpretation makes Korean a Differential Subject Marking (DSM) language where the subject has different formal coding depending on its semantic, pragmatic or other properties.

Whereas interpretive differences correlating with DSM in Korean have been extensively studied in the literature (Ahn & Cho 2006a, 2006b, 2007, Kwon

& Zribi-Hertz 2008, Chung 2015, H. Lee 2010, 2016, 2017, Park 2020, among others), the association of caseless-subject clauses with a direct perception interpretation exemplified in (1b) has not been systematically explored on the basis of a close examination of naturally occurring data. The question to be explored in the present study is therefore: Why is it that the caseless-subject pattern is the only available option to trigger an interpretation based on the speaker's direct perception of an event in the here and now?

The main purpose of this paper is to provide a novel account of DSM that can explain the association of the absence of case marking with direct perception in the here and now. Based on evidence from analyses of conversation data, I will argue that productive occurrence of caseless subjects in eventive direct perception clauses follows from an economical use of formal particles motivated by cue reliability, that is, reliability of linguistic and situational cues for predicting and identifying an argument function.

The remainder of this paper is structured as follows. Section 2 provides a review of theoretical approaches to DSM in Korean. Section 3 presents evidence from conversation data demonstrating that caseless subjects most productively occur in clause types with a subject that is identifiable in the here and now as an agent directly by the speaker and/or the hearer. In Section 4, I extend Lestrade & de Hoop's (2016) analysis of tense/aspect-based DSM in terms of grounding to account for patterns of uses of subject marking in Korean, arguing that the association of the absence of case marking with direct perception of an event follows from an economical use of case marking that is made possible by grounding. Going one step further, I propose a uniform explanation for the association of caseless subjects with direct perception in the here and now, agentivity and tense deficiency in terms of cue reliability (Levshina 2021). If there are strong and reliable cues for identifying an argument function and tense available to the speaker and/or the hearer, the use of formal particles indicating such properties can be judged redundant and suspended because of economy. Section 5 concludes the paper by discussing limitations and implications of the present study.

2. Previous Accounts of DSM in Korean

The term ‘caseless’ subjects is used in this study to refer to those that occur without functional particles signaling case or discourse function. Caseless nominals (Ahn & Cho 2006a, 2006b, 2007) have been noticed and discussed for Japanese and Korean under various labels, for example, bare nominals (Kwon & Zribi-Hertz 2008), case particle ellipsis (Tsutsui 1984, Watanabe 1986, Fry 2003, H. Lee 2015, 2016), case deletion or drop (Masunaga 1988, Ahn 1999, Minashima 2001, S. Lee 2006, Chung 2015), and zero particle (D. Lee 2002, Shimojo 2005, Kim 2008). This section examines previous accounts of caseless subjects. I will group these accounts under two headings, namely (i) information structure-based and (ii) syntax-based accounts, and review two accounts which exemplify the two broad classes of analyses of Korean DSM in detail.

2.1. Kwon & Zribi-Hertz’s (2008) Information Structure-based Account

Kwon & Zribi-Hertz (2008) have proposed an account of Differential Marking (DM) that captures the interpretive contrasts between case-marked arguments and caseless arguments in terms of f(ocus)-structure. They attempt to derive a range of distributional and interpretive differences between case-marked subjects and caseless subjects from a single information structural parameter—f-structure visibility. Kwon & Zribi-Hertz’s (2008) proposal is summarized in (2):

- (2) DM and f-structure in Korean (Kwon & Zribi-Hertz 2008: 279)
 - a. NPs that support functional markers indicating structural positions in syntax are visible in f-structure.
 - b. NPs that fail to support such markers are not visible in f-structure, unless some other type of marking guarantees their visibility as f-structure constituents.

The term f-structure, as used by Erteschik-Shir’s (1997, 2007), identifies a level of grammatical representation where the output of syntax is annotated

for information packaging. As does Lambrecht (1994), Erteschik-Shir (1997, 2007) regards f-structure as a structural description where foci, instantiating new information, are paired up with topics, instantiating presupposed or old information. Under her theory, f-structure constituents are either topics or foci. Hence, what (2) means is that when overt subjects or objects fail to support a functional marker in morphology, they cannot be identified as topics or foci at any level of f-structure.

Adopting Erteschik-Shir's (1997, 2007) framework to represent f-structure, Kwon & Zribi-Hertz (2008) argue that like caseless objects, caseless subjects abide by (2b), in that they fail to be visible in f-structure and consequently undergo f-structure incorporation. Under this analysis, the interpretive contrasts between caseless-subject clauses and *-(n)un/-ka*-marked subject clauses follow from different f-structure properties. In (3) and (4) below, I illustrate f-structure patterns of DSM by the examples in (1). The f-structure representations Kwon & Zribi-Hertz (2008) propose for the two readings of (1a) are given in (3).¹

- (3) Pesu-ka o-n-ta.
 bus-Nom come-Pres-Decl
 'There is the/a bus coming.' or 'It is the/a bus that is coming.'
 a. (What is going on?/Why am I hearing strange noises?)
 [∅]s.TOP1 [[the bus]TOP2 [is coming]FOC2]FOC1
 b. (Is the taxi coming? No.)
 [<the bus>FOC]TOP is coming

(3a) is a complex f-structure involving two levels: the matrix level (level 1) and the embedded level (level 2). This f-structure pattern typically responds to such questions as *What's up?* or *What is going on?*. The reply is event-reporting or presentational; hence it is spatiotemporally anchored and hosts a 'stage-level' predicate. Kwon & Zribi-Hertz (2008) assume that thethetic subject is included within the new information—the matrix focus (FOC1). Lambrecht (1994) analyzes

¹ Following Kwon & Zribi-Hertz's (2008) notation to represent f-structure, category labels placed outside closing brackets indicate f-structure constituents, while labels placed inside opening brackets identify s-structure. When a sentence involves two or more levels of f-structure, I use digits to help the reader associate each focus with the appropriate topic, e.g.: [.....]TOP1 [[.....]TOP2 [.....]FOC2]FOC1

such sentences as containing no topic. Erteschik-Shir, by contrast, assumes that such clauses contain a stage topic denoting a spatiotemporal discourse referent—the time and place to which the reported event or situation is anchored. Kwon & Zribi-Hertz (2008) further assume that a thetic (event-reporting or presentationl) clause such as (1a) has an embedded f-structure correlating with the predication relation between the subject and predicate.

Kwon & Zribi-Hertz (2008) propose to analyze caseless-subject clauses as having no internal f-structure, as illustrated in the proposed f-structure in (4) for the example in (1b). Thus under this analysis, the *-ka*-subject thetic clause in (1a) and the caseless-subject clause in (1b) differ with respect to the visibility of the syntactic subject in f-structure.

- (4) Pesu o-n-ta.
 bus come-Pres-Decl
 '(Watching coming of the bus) Here comes the bus.'
 [Ø]s.TOP [the bus comes]FOC

Kwon & Zribi-Hertz (2008) contend that caseless subjects, like caseless objects, can be construed as neither active topics nor as foci, and always occur in tense-deficient clauses construed as thetic and anchored to speech time. They account for the correlation between the simple thetic interpretation and the tense deficiency of caseless-subject clauses in terms of f-structure invisibility.² Since they are invisible in f-structure, caseless subjects must be included within the matrix focus of their clause, and hence must partake in a thetic interpretation. Unlike nominative-marked thetic subjects, however, caseless subjects are invisible in f-structure, and hence do not stand as embedded topics; it follows that caseless subjects must be incorporated into the matrix focus of their thetic f-structure. Kwon & Zribi-Hertz (2008) further claim that caseless-subject clauses may be specified for aspect, but are left unspecified for tense. Since they are construed as thetic clauses and must consequently be temporally anchored, caseless-subject clauses involve pragmatic anchoring to speech time,

² Tense-deficient clauses are used here to refer to clauses that are left unspecified for tense, although they may be specified for aspect. Section 4.2 will provide a more concrete characterization of tense-deficient vs. tense-specified clauses with examples.

as in (1a), whereas topical subject and nominative-markedthetic subjects fail to be similarly restricted.

Kwon & Zribi-Hertz's (2008) information structure-based account is theoretically attractive in that it derives a range of observed properties of DM from a single f-structure property of arguments, i.e., f-structure visibility: it technically accounts for the observed interpretive and distributional properties correlating with DSM in terms of f-structure visibility. Unfortunately, the f-structure invisibility of caseless subjects remains unmotivated in Kwon & Zribi-Hertz's (2008) account.³ It is unclear (i) why caseless subjects are restricted to tense-deficient clauses anchored to speech time? and (ii) why caseless-subjectthetic clauses trigger a direct perception interpretation as in (1b) whereas *-ka*-subjectthetic clauses do not?

2.2. Park's (2020) Syntax-based Account

Park (2020) proposes a novel account of DSM that captures the interpretive contrast between case-marked and caseless subjects in terms of the two types of definiteness distinguished by Schwarz (2009). The relevant definiteness contrast correlating with DSM is summarized in (5) and illustrated in narrative sequences in (6).

(5) Subject forms and types of definite reading

Subject form	Definiteness
<i>-(n)un/-ka</i> -marked, bare noun	Anaphoric (strong) definiteness
Demonstrative + noun	
Unmarked, bare noun	Uniqueness (weak) definiteness

³ See Chung (2015), H. Lee (2015, 2017) and Park (2020) for critical reviews of theoretical and empirical problems of Kwon & Zribi-Hertz (2008).

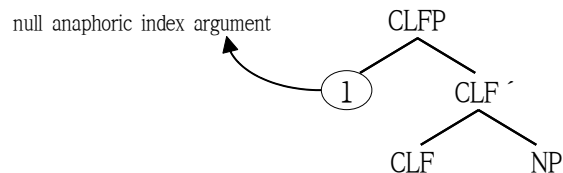
- (6) Korean narrative sequences: Bare nouns and demonstratives in subject positions (Park 2020: 11)

[Haksayng₁-i] ecey pak kyoswu-lul chacawa-ss-ta.
 student-Nom yesterday Park professor-Acc visit-Pst-Decl
 (Kuliko) [(ku) haksayng₁-(i,un)]/ haksayng₁-(i,un)]
 (and) that student-{Nom,Top}/ student-{Nom,Top}
 pak kyoswu-eykey chwuchense-lul pwuthakhay-ss-ta.
 Park professor-Dat recommendation-Acc ask-Pst-Decl
 'There was a student who visited Prof. Park. {(That) student} asked Prof. Park to write a recommendation letter for him.'

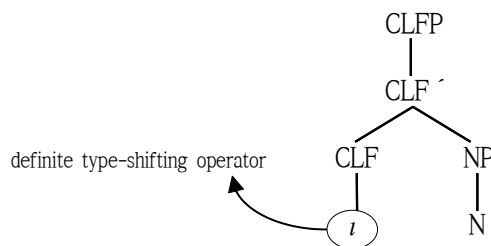
Here, the first sentence introduces a novel discourse referent (a student), and the second sentence must refer back to this referent. According to Park (2020), the subject of the second sentence in this narrative sequence can be construed as an anaphoric definite referring back to *haksayng* 'student' introduced in the first sentence. By contrast, the caseless, bare (unmodified) subject in the second sentence cannot refer back to *haksayng* 'student' in the first sentence.

Park (2020) proposes that this difference follows from syntactic differences between marked and unmarked (caseless) bare common NPs in Korean shown in (7) and (8). His analysis builds on the idea that anaphoric definites carry a strong article whose anaphoric index argument establishes an anaphoric link to an explicit linguistic antecedent (Schwarz 2009, Jenks 2018, Park & Kang 2020, among others). Building on this idea, Park (2020) proposes that the null anaphoric index argument is structurally accommodated at the periphery of CLFP (Classifier Phrase), which is a bare NP in articleless languages like Korean, and is lexicalized by case morphology as in (7). He further argues that caseless, bare subject NPs achieve definite interpretations via a type-shifting operator optionally inserted to the CIF (Classifier) head position, as in (8).

- (7) *-(n)un/-ka*-marked, bare subjects as anaphoric definites



(8) Unmarked, bare subjects as unique definites



An important consequence of this analysis is that caseless subject NPs, unless preceded by a demonstrative, cannot be construed as an anaphoric definite, thus correctly predicting the infelicity of a caseless-subject clause in anaphoric environments, as illustrated in (6).

A further consequence of this analysis is that while anaphoric definite bare subject NPs must be overtly marked, unique definite subject NPs do not require case marking, as shown in (9).

- (9) a. Elevator(-ka) ecey kocangna iss-ess-ta.
 elevator-Nom yesterday break.down be-Pst-Decl
 'The elevator broke down yesterday and was not working.'
- b. Radio(-ka) sikkulep-ney.
 radio-Nom be.noisy-Excla
 'The radio is noisy.'
- c. TV(-ka) khye-iss-ta.
 TV-Nom be.on-Decl
 'The TV is on.'
- (Park 2020: 13)

Thus, Park's analysis accounts for the general optionality of case marking on subjects which occur in unique definite environments. As predicted by this analysis, the caseless subject form is generally not obligatory in stative direct perception clauses, as shown in (9). By contrast, the two subject forms are not equally acceptable in eventive direct perception clauses, as illustrated in (10). Whereas both the case-marked and caseless subject forms are acceptable in (9), the caseless form is the better option in (10). However, Park's analysis is concerned with the nominal structure, i.e., the structure of differentially

marked bare common nouns, and its interface with their referential properties. Hence it does not account for the contrast between (9) and (10) that holds at the clause or utterance level, failing to explain why the case-marked form is not felicitous in eventive direct perception clauses describing an agent's action.

- (10) a. Ne/*ney-ka mwe mek-ni?
 you/you-Nom what eat-Q
 'What are you eating?'
 b. Ceki Minswu/?Minswu-ka nao-n-ta.
 there Minsoo/Minsoo-Nom come.out-Pres-Decl
 'Minsoo is coming out over there.'

Another challenge for Park's (2020) account is the fact that as mentioned in Section 1, only the caseless subject pattern triggers a direct perception interpretation in both eventive and stative clauses. For instance, the caseless-subject clause in (9b) is preferably construed as a statement about the noise that is sensed directly by the speaker at the moment of utterance, where as the case-marked counterpart is understood as a statement of a noise that is not necessarily based on the speaker's direct perception of the situation. Despite the considerable amount of work on Korean DCM, little work has asked (i) why the caseless subject pattern is more strongly preferred in eventive direct perception clauses compared to in stative direct perception clauses, and (ii) why only this pattern triggers a direct perception interpretation.

In this paper, I propose an alternative analysis of Korean DSM in terms of economy motivated by cue reliability that can account for hitherto unexplained contrasts between caseless- and case-marked subject clauses discussed in this section. Before discussing this proposal, I first present in Section 3 an analysis of whether and how caseless-subject clauses correlate with direct perception in conversation data.

3. Grounding and Caseless Subjects: Evidence from Conversation Data

This section presents an analysis of conversation data that aims to explore empirical evidence for the correlation between the absence of case marking on subjects and direct perception of an event in the here and now.

3.1. The Data

The conversation data for this study come from sixty-seven hours of video-taped conversation between four pairs of native speakers of Korean, who were born and raised in Korea, and enrolled in the university at the time of the recording. The paired participants agreed to have their conversation video-taped in their apartments for periods ranging from four to ten days in July 2020. The participants – four females and four males – were between the ages of 22 and 28, and were mutual friends. The casual nature of the conversation is clearly indicated by the predicate form which they used in their conversation; they predominantly used the plain (i.e., casual) form of predicates consistently, except for one pair, in which the distal (i.e., polite) forms were mixed with the plain forms.

The recorded conversations were mostly conversations among the eight participants who are mutual friends though they occasionally include conversations with other speakers. These conversations were transcribed by three Korean graduate students who majored in linguistics and prepared in terms of clausal units. Utterance boundaries were identified according to pause and the transcribed text was divided into units accordingly. Pause could be identified fairly clearly and was found to be consistent index to identify utterance boundaries. However, a pause caused morphosyntactically unnatural divisions occasionally, for example, dividing a sentence final particle and the predicate to which it is attached, and in these cases the pause was ignored. After utterance units were identified by pause, they were further divided into clausal units. Complex sentences are divided into clausal units, regardless of subordination types. Hence, an adverbial subordinate clause, a noun complement clause, and

a nominalized clause were all considered as separate clausal units.

The procedures outlined above produced about 1000–2000 clausal units for each conversation pair, and 7780 clausal units for all pairs in total, which include 1396 tokens of caseless-subject clauses and 1516 tokens of nominative-subject clauses. Of these clauses, 1250 tokens of caseless-subject clauses were analyzable in the sense that they were clearly referential (i.e., not grammatical items, question words, etc.) and bore well-defined semantic roles. These clausal units were manually coded for clause type distinguished according to the degree of grounding by the same graduate students who transcribed the recorded conversations and then checked by the author.

3.2. The Notion of Grounding and Classification of Clause Types

In their recent work on DSM in Hindi, Nepali and Manipuri, Lestrade & de Hoop (2016) convincingly demonstrate that grounding plays a crucial role in motivating the tense/aspect-based differential use of ergative case and the differential use of ergative case driven by the distinction between stage- and individual-level predication. According to Dixon (1979) and DeLancey (1981), events in the present or imperfect are “A(gent)-centered” in the sense that these events have an identifiable agent: the agent function of an event participant of ongoing events in principle can be identified by the hearer on the basis of non-linguistic information available in the here and now.⁴ Building on this insight, Lestrade & de Hoop (2016) propose that the agent function of an event participant of an ongoing event can be *grounded* in the here and now (Lestrade & de Hoop 2016: 403). When the activity is finished, however, the hearer can no longer see the agent at work. As such, Lestrade & de Hoop (2016) assume that the agent function of an event participant of past events cannot be grounded in the here and now.

The notion of grounding originally proposed by these authors to account for tense/aspect-based differential use of ergative case is defined as the possibility of *the hearer* to determine an agent function of an event participant

⁴ We understand the notion of agent in a broad sense here, also applying to experiencers and entities capable of moving themselves along the lines of Dowty’s (1991) theory of proto-roles.

himself (Lestrade & de Hoop 2016: 398). Their analysis based on this notion of grounding would predict that the agent of an event be marked with case when the hearer cannot in principle evidence the event directly. But this prediction is not borne out in Korean, as evidenced by the preference for the caseless-subject clause in (11).⁵

- (11) Pesu o-n-ta. / #Pesu-ka o-n-ta.
 bus come-Pres-Decl/ bus-Nom come-Pres-Decl
 (The speaker is looking at coming of the bus at the bus stop and talks to the hearer over the phone.)
 'Here comes the bus.'



Here only the speaker can see coming of the bus. As such, the role played by the subject's referent in the event is not directly identifiable by the hearer who does not share the here and now with the speaker. For such a case,

⁵ An anonymous reviewer points out that the contrast between the caseless- and case-marked subject clause seems to disappear when the exclamative speech act sentence particle *-ney* is used, as illustrated in (i):

- (i) Pesu o-ney. / Pesu-ka o-ney.
 bus come-Excla bus-Nomcome-Excla
 'Here comes the bus!'

It is an interesting observation that the nominative-marked subject NP felicitously occurs in eventive direct perception clauses in which the verb stem is followed by the exclamative speech act suffix *-ney*, indicating a surprise. How and why the use of different speech act particles influences the naturalness of case marking on subjects is an empirical question which requires a more thorough investigation in future study.

an explanation in terms of the hearer-oriented notion of grounding wrongly predicts the subject be marked with case, contrary to the fact.

In this paper, I suggest that Lestrade & de Hoop's (2016) analysis of differential use of ergative case can straightforwardly extend to DSM in Korean if the notion grounding is broadened to incorporate the speaker's role in deriving the argument structure as well and understood as the possibility of the speaker *or* the hearer to determine the argument function of an event participant themselves on the basis of situational information available to them in the here and now. I further suggest that grounding should be viewed as a gradable property that depends on the extent to which the speaker and/or the hearer can directly determine the argument function of an event participant because determination of an argument function can be grounded in the here and now in different degrees. The table in (12) shows a classification of clause types according to degrees of grounding. As shown in the table, clauses in the data can be classified into two broad types according to whether or not identification of the agent function of an event participant can be grounded in the here and now: clauses with an agent that can be grounded in the here and now and clauses with an agent that cannot be grounded in the here and now. I will refer to the former type as [+Grounded] clauses and the latter type as [-Grounded] clauses.

(12) Classification of clause types

Clause type		Determination of argument function
[+Grounded]	Type 1	Directly identifiable by the speaker and the hearer in the here and now
	Type 2	Directly identifiable by the speaker or the hearer in the here and now
[-Grounded]	Type 3	Determined on the basis of direct perception of the completed action or direct perception of the result of the completed action
	Type 4	Determined on the basis of other sources of information

Crucially, the transition between [+Grounded] and [-Grounded] clauses is gradual because determination of an argument function can be grounded in the here and now in different degrees. As shown in the above table, both [+Grounded] and [-Grounded] clauses can be divided into two subtypes according to degrees of grounding. This finer-grained classification of clauses into four subtypes is discussed in detail in the following subsection with an analysis of the proportion of these subtypes.

3.3. Analyzing the Distribution of Subtypes of Caseless-Subject Clauses

Type 1 clauses describe the state of affairs that can be directly observed by both the speaker and the hearer. An example of Type 1 clause in the data is given in (13).

- (13) (Speaker A and Speaker B are standing by the window in the living room waiting for food delivery. Speaker A utters the following, while looking at approaching of the delivery motorcycle to the entrance of their apartment.)
 Paytal accessi o-n-ta!
 delivery man come-Pres-Decl
 'Here comes the delivery man!'

Here the speaker and the hearer share the here and now, and talk about an entity within the context.

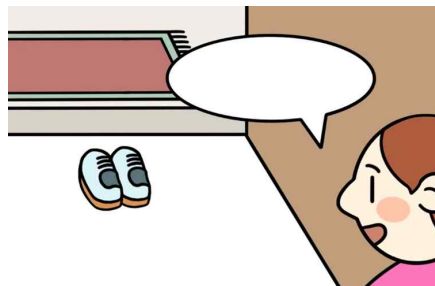
Type 2 clauses differ from Type 1 in that the speaker and the hearer do not share the here and now. There are two subtypes of Type 2 clauses. The first type either expresses the speaker's subjective perceptions (e.g., feeling of hunger, loneliness, fatigue, etc.), or describes situations that the speaker is directly observing. Examples in (1b) and (11) above exemplify this subtype, which describes a situation that is directly sensed only by the speaker. The second subtype of Type 2 clauses in the data are statements or questions about situations that the speaker believes that the hearer can check on the basis of information available to her in the here and now. An example of this subtype is exemplified by Speaker A's utterance in (14).

- (14) (Speaker A has an appointment with a CEO of some company. He arrives at the CEO's office and asks Speaker B, his secretary about the CEO's presence in his office.)

A: Kim tayphyo-nim an-ey kyeysi-eyo?
 Kim CEO-Hon inside-Loc be.Hon-Pol
 'Is President Kim in?'
 B: Ney. Camkkan kitalli-si-eyo.
 Yes a.second wait-Hon-Pol
 'Yes, please wait a second.'

Here Speaker B may not be directly observing the agent, president Kim, at the moment of Speaker A's utterance, but Speaker A assumes that Speaker B in principle can check directly whether president Kim is in his office or not. As such, the caseless subject-clause in Speaker A's utterance in (14) may be regarded as reflecting the speaker's belief about the hearer's ability to directly determine the argument function. Type 3 clauses describe finished activities and have no identifiable agent as the speaker can no longer see an agent at work. Nevertheless, they involve direct perceptions: they are based on the speaker's direct perception of the result of the agent's action, as exemplified by (15). Here the speaker utters the sentence upon witnessing the shoes left at the entrance of her house that she thinks are Youngmi's shoes, which can be taken as the evidence suggesting her arrival.

- (15) Yengmi o-ss-ta!
 Youngmi come-Pst-Decl
 'Youngmi has come!'



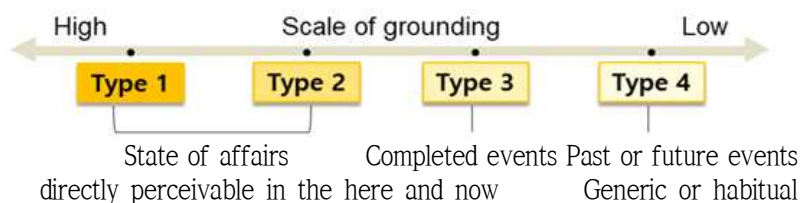
By contrast, Type 4 clauses do not involve direct perceptions: they are either generic or habitual clauses or clauses stating or asking about future events. An example of Type 4 clause is given in (16) (Speaker B's utterance).

- (16) A: Ne yocum way kulehkey pappā?
 you these.days why so be.busy
 'How come you are so busy these days?'
 B: Nayil mikwuk-eyse wuli enni wa.
 tomorrow U.S.-from our sister come
 'My sister is coming from the U.S. tomorrow.'
 Kulayse kathi yehayng ka-l cwunpi hay.
 so together trip go-Rel preparation do
 'So I'm preparing for going on a trip with her.'

Unlike in examples given above, Speaker B's utterance in (16) describes an event that is still to happen, i.e., her sister's visit, and cannot be construed as based on the speaker's direct observation.

Classification of clause types according to the degree of grounding is summarized in (17).

(17) Classification of clause type according to degrees of grounding

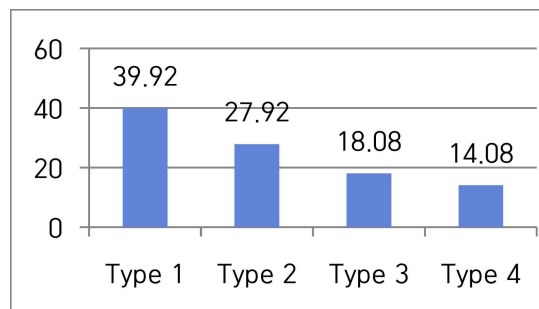


Here the scale of grounding can be conceived of as derived from the degree to which the agent function of an event participant can be determined on the basis of information available from the here and now. Clause types higher on this have a subject which is most straightforwardly identifiable as an agent of an ongoing event. On the other hand, clause types lower on the scale denote events that have been completed or took place in the past or future events. So, in such clauses, the agent is no longer a viewpoint focus and is no longer

be directly identifiable by the speaker or the hearer in the here and now.

Using the system of analysis summarized in (12), I analyzed clauses with caseless subjects in the data in terms of four categories, leaving an analysis of nominative-marked subjects to future work. (18) shows the proportion of the four subtypes of caseless-subject clauses.

(18) Proportion of four subtypes of caseless-subject clauses



As can be seen, the relative frequency of clauses with caseless subjects decreases steadily from left and right: Type 1 was the most frequently produced subtype (39.92%), and Type 4 was the least frequent subtype (14.48%). We can also observe that [+Grounded] clauses (Type 1 and Type 2 clauses) showed a higher rate compared to [-Grounded] clauses (Type 3 and Type 4): 67.84% of all tokens of caseless-subject clauses were [+Grounded] clauses.

To summarize, the results of the frequency analysis of subtypes of caseless-subject clauses indicate that caseless subjects most productively occur in clause types with higher degrees of grounding, thus providing empirical evidence for the association between the absence of case marking on subjects and high degrees of grounding.

4. Accounting for Properties of Caseless-Subject Clauses

This section proposes a uniform explanation for the association of caseless subjects with direct perception in the here and now, and with other factors such as agentivity and tense deficiency in terms of an economical use of formal

particles motivated by cue reliability.

4.1. Direct Perception, Cue Reliability and Caseless Subjects

The results of the analysis of the conversation data discussed in Section 3.3 raises the following question: Why is it that caseless subjects most productively occur in the clause types expressing direct perception of a state of affairs by the speaker and/or the hearer in the here and now? Building on Lestrade & de Hoop (2016), I argue that the association of the absence of case marking on subjects with direct perception utterances follows from an efficient use of case marking that is made possible by grounding: The use of case marking on subjects becomes more redundant in clause types higher on the scale of grounding because there are increased clues to the agent of an event in context, i.e., the here and now, as one moves up the scale. Following a general economy principle, the speaker can omit the explicit use of case marking to minimize her effort if the here and now can be used by the speaker and/or the hearer to derive the argument structure. By contrast, higher degree of explicitness in subject marking is more necessary in clause types lower on the scale of grounding in order to indicate to the hearer that the agent is no longer straightforwardly identifiable in the context.

Under this analysis, the productive occurrence of the caseless subject form in direct perception clauses can be explained as resulting from conventionalization of the speaker's preference to omit redundant case marking. The ability of caseless-subject clauses to trigger a direct perception interpretation can be seen as an interpretive correlate of their probabilistic property in language use.

What remains to be explained is why the case-marked and caseless-subject forms are not equally preferred in eventive and stative direct perception clauses. As noted in Section 2.2, both forms are generally acceptable in stative direct perception clauses whereas the caseless form is strongly preferred in eventive direct perception clauses. This difference between eventive and stative direct perception clauses is evidenced by the contrast between (9) and (10), repeated here as (19) and (20), respectively.

- (19) a. Elevator(-ka) ecey kocangna iss-ess-ta.
 elevator-Nom yesterday break.down be-Pst-Decl
 'The elevator broke down yesterday and was not working.'
- b. Radio(-ka) sikkulep-ney.
 radio-Nom be.noisy-Excla
 'The radio is noisy.'
- c. TV(-ka) khye-iss-ta.
 TV-Nom be.on-Decl
 'The TV is on.' (Park 2020: 13)
- (20) a. Ne/*ney-ka mwe mek-ni?
 you/you-Nom what eat-Q
 'What are you eating?'
- b. Ceki Minswu/??Minswu-ka nao-n-ta.
 there Minsoo/Minsoo-Nom come.out-Pres-Decl
 'Minsoo is coming out over there.'

The strong preference for the caseless subject form in eventive direct perception clauses may be accounted for in terms of an efficient use of case marking motivated by cue reliability, that is, reliability of informational cues for predicting a particular syntactic role or grammatical function. Cue reliability, which is sometimes also called cue validity (Rosch & Mervis 1975), is high when the cue is never misleading or ambiguous and always leads to correct expectations. It can be stated as the probability of a grammatical function given a cue, and can be expressed as $\text{Pr}(\text{obability}) (\text{Role} \mid \text{Cue})$ using the notation from probability theory.

Cue reliability shows how much the hearer can rely on informational cues such as particular formal, semantic, or pragmatic features of an argument or contextual information in order to predict a particular syntactic role. As pointed out by Levshina (2021), it is important for efficiency-based accounts of communicative behavior which assume that the speaker minimizes his or her effort, while at the same time making sure that the hearer interprets correctly who did what to whom. If the cues are reliable and strong enough to guide the hearer to decide on the correct syntactic role of an argument, then the speaker can omit case marking because additional marking is redundant. If they are weak or likely to lead to wrong predictions, then the speaker is more likely to provide additional cues such as the case marker in order to

help the hearer to correctly identify the syntactic role of an argument. Over time, this efficient behavior may become conventionized and grammaticalized (Hawkins 2004, Haspelmath 2008, Jaeger 2010, Levshina 2021).

In many languages, semantic properties of arguments such as agent properties and patient properties serve as important cues for predicting a subject role and an object role, respectively. As discussed in Section 3.2, eventive [+Grounded] clauses have a subject that is most straightforwardly identifiable as the agent of ongoing events. High agentivity serves as a reliable cue for predicting and identifying the subject in the clause because prototypical agents occur more frequently as subjects than as objects. Thus, in eventive [+Grounded] clauses, there are two kinds of reliable and strong cues that can guide the hearer to identify an agent, i.e., the semantic property of an argument (agentivity) and direct evidence from the context (situational information from the here and now). The availability of such reliable cues makes additional marking on subjects redundant, leading case marking be more easily omitted by the speaker in eventive [+Grounded] clauses. This explains why the caseless subject form is more strongly preferred to the case-marked form in eventive [+Grounded] clauses.

By contrast, subjects of stative [+Grounded] clauses are either less prototypical as an agent or lack agent properties at all. Thus in a stative context the semantic feature of the subject referent does not serve as a reliable cue for a subject role though the contextual cue may still be available to the speaker or the hearer. If cue reliability influences the probability of subject marking, we can expect that the speaker is more likely to provide additional cues in situations of low cue reliability in order to help the hearer to correctly identify the argument role. This account of DSM based on cue reliability predicts that case marking on subjects will be less likely to be omitted in stative [+Grounded] clauses compared to in eventive [+Grounded] clauses. Whether this prediction is indeed confirmed by an analysis of naturally occurring data I leave for future work.

4.2. Tense Deficiency and Caseless Subjects

A further consequence of this analysis is that it is able to offer a uniform motivation for seemingly unrelated properties of caseless-subject clauses—high

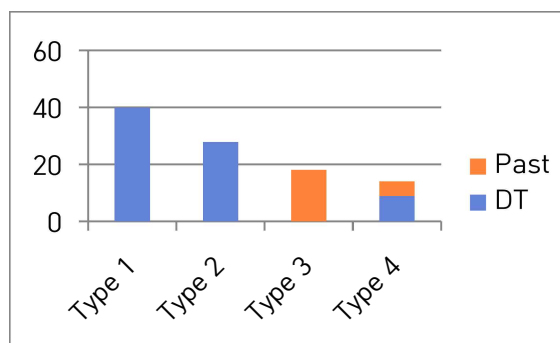
degrees of grounding and agentivity and tense deficiency. Case and tense seem unrelated at first sight as the first one pertains to nominal domain, denoting properties of individuals, whereas the second one pertains to the verbal domain, denoting properties of events. Case alternations triggered by tense and aspect, however, are not at all uncommon in and across languages (See Malchukov & de Hoop (2011) for a review of patterns of tense/aspect and mood based case alternations).

The association of caseless subjects with tense deficiency was examined by comparing the distribution of caseless-subject clauses in deficient tense and past tense conditions. In the deficient tense (DT) condition, clauses were syntactically unspecified for tense. Such tense-deficient clauses include sentences which host no tense or aspect marker, as in examples in (20a), and clauses which host markers signalling aspect but not tense, as in (20b). As pointed out by Kwon & Zribi-Hertz (2008) and many others, the affix *-(nu)n*, commonly glossed as ‘present’ tense, crucially does not signal temporal anchoring to speech time: like the English simple present, it occurs in generic and habitual clauses and may also indicate future, for instance. As such, this affix has been taken as an aspect marker unspecified for tense by some scholars (e.g., Kwon & Zribi-Hertz 2008) or as a non-past marker by others (e.g., E.H. Lee 2019).

In the past tense condition, clauses are specified for the affix *-(e/a)ss*, commonly glossed as past tense. In addition to past, this affix may also indicate completion, as illustrated in (15), which implies that the resulting state still obtains at the utterance time. Following E.H. Lee (2019: 98–99), I will take *-(e/a)ss* as the past-tense marker and assume that its apparent aspectual meaning, that is, the meaning of completion and current relevance, comes from its status as a relative past marker and from the lexical aspectual feature of the verb.

Let us now look at the proportion of DT and past tense caseless-subject clauses. In our Korean data, a majority of attested caseless subjects occur in tenseless clauses anchored to speech time. 76.72% of all tokens of caseless-subject clauses were found in the DT condition, whereas 23.28% were found in the past tense condition. Examining the proportion of DT and past tense caseless-subject clauses within each of the four subtype analyzed in Section 3.3, we find the results summarized in (21).

(21) Proportion of DT and past tense clauses within four subtypes



As can be seen, all tokens of Type 1 and Type 2 clauses were tense-deficient, whereas all tokens of Type 3 clauses were found in the past tense condition. The figure in (21) further shows that 36.93% of Type 4 clauses were found in the past tense condition. Thus, these results strongly suggest that the correlation between the absence of case marking on subjects and tense deficiency is not absolute but a matter of statistical preferences.

Under the analysis of DSM I propose here, the two salient properties of caseless-subject clauses—direct perception in the here and now and tense deficiency—both follow from an economical use of formal particles motivated by cue reliability. These two properties may be tied up in the following way. As noted in Section 3.3, a salient pragmatic feature of a majority of attested caseless-subject clauses in the data is that the speaker and/or the hearer can ground the agent function of an event in the here and now. Because situational information available from the here and now serves as a strong cue for identifying the argument function of an event participant, the speaker may suspend the use of case marking for economy reasons.

When the speaker and/or the hearer is able to sense an event directly in the here and now, they can ground not only argument functions but also temporal interpretations of the event in the here and now. For instance, when they observe an ongoing event directly at the moment of utterance, they can identify the time at which an action is happening directly in the here and now. Hence the speaker can be more economical in his or her utterance, omitting both case and present tense markers. Due to this tense deficiency, caseless-subject clauses are interpreted as being pragmatically anchored to utterance time as a default option.

To summarize, this section has proposed a uniform explanation for the association of caseless subjects with direct perception in the here and now, agentivity and tense deficiency in terms of cue reliability. If there are strong and reliable cues for identifying an argument function and tense available to the speaker and/or the hearer, the use of formal particles indicating such properties can be judged redundant and suspended because of economy.

5. Conclusion

Caseless and case-marked arguments in Korean have been shown to differ systematically as to their interpretation. This paper has focused on a hitherto unexplained difference between caseless- and case-marked subject clauses in the availability of a direct perception interpretation. Evidence from conversation data demonstrates that caseless subjects most productively occur in clause types with a subject that is identifiable in the here and now as an agent directly by the speaker and/or the hearer. Based on this evidence, I have argued that the association of the absence of case marking on subjects with direct perception clauses follows from an efficient use of case marking that is made possible by grounding. Going one step further, I have proposed a uniform explanation for the association of caseless subjects with direct perception in the here and now, agentivity and tense deficiency in terms of cue reliability. Thus, this analysis shows that it is possible to develop a unifying account of such seemingly unrelated properties of caseless-subject clauses which subsumes them under the single principle of economy motivated by cue reliability. These preliminary results support efficiency-based accounts of case marking and grammar (Hawkins 2004, Haspelmath 2008, Jaeger 2010, H. Lee 2010, 2016, Lestrade & de Hoop 2016, Levshina 2021) and underscore the importance of communicative efficiency in explaining and motivating patterns of grammar and language use.

Nevertheless, this study has an important limitation in its scope in that it analyzed caseless-subject clauses only and did not compare them to case-marked subject clauses. Further research is needed to investigate how case-marked subject clauses pattern differently from caseless- subject clauses in naturally

occurring data with respect to factors analyzed in this study.

Final open questions are whether and how cue reliability influences the probability of object marking as well and whether ellipsis in general can be explained by an economical use of linguistic elements (encompassing both lexical elements and formal particles) motivated by cue reliability. Investigation of these issues would provide a fruitful avenue for future research.

References

- Ahn, Hee-Don. 1999. "Notes on Case Deletion," in Yong-Wha Kim, Il-Kon Kim, and Jeong-Woon Park, eds., *Linguistic Investigation: In Honor of In-Seok Yang*, 1-16. Seoul: Hanshin Publishing Co.
- Ahn, Hee-Don and Sungeun Cho. 2006a. "Wh-topics and Unpronounced Cases on Wh-phrases," *Studies in Generative Grammar* 16(1): 61-90.
- Ahn, Hee-Don and Sungeun Cho. 2006b. "Layered Nominal Structures: Implications for Caseless Nominals," *Korean Journal of Linguistics* 31(2): 165-185.
- Ahn, Hee-Don and Sungeun Cho. 2007. "Subject-Object Asymmetries of Morphological Case Realization," *Language and Information* 11(1): 53-76.
- Aissen, Judith. 2003. "Differential Object Marking: Iconicity vs. Economy," *Natural Language and Linguistic Theory* 21(3): 435-483.
- Bossong, Georg. 1985. *Differentielle Objekmarkierung in der Neuiranischen Sprachen*. Tübingen: Gunter Narr Verlag.
- Chung, Eun Seon. 2015. "Challenging a Single-Factor Analysis of Case Drop in Korean," *Language and Information* 19(1): 1-18.
- DeLancey, Scott. 1981. "An Interpretation of Split Ergativity and Related Patterns," *Language* 57(3): 626-657.
- Dixon, R. M. W. 1979. "Ergativity," *Language* 55(1): 59-138.
- Dowty, David. 1991. "Thematic Proto-roles and Argument Selection," *Language* 67(3): 547-619.
- Erteschik-Shir, Nomi. 1997. *The Dynamics of Focus Structure*. Cambridge: Cambridge University Press.
- Erteschik-Shir, Nomi. 2007. *The Syntax/Discourse Interface: Information Structure*. New York: Oxford University Press.
- Fry, John. 2003. *Ellipsis and Wa-marking in Japanese Conversation*. New York: Routledge.
- Haspelmath, Martin. 2008. "Frequency vs. Iconicity in Explaining Grammatical Asymmetries," *Cognitive Linguistics* 19(1), 1-33.

- Hawkins, John. 2004. *Efficiency and Complexity in Grammar*. Oxford: Oxford University Press.
- de Hoop, Helen and Peter de Swart (eds.). 2008. *Differential Subject Marking*. Dordrecht: Springer.
- Jaeger, T. Florian. 2010. "Redundancy and Reduction: Speakers Manage Syntactic Information Density," *Cognitive Psychology* 61(1): 23-61.
- Jenks, Peter. 2018. "Articulated Definiteness without Articles," *Linguistic Inquiry* 49(3): 501-536.
- Kim, Taeho. 2008. *Subject and Object Markings in Conversational Korean*. Buffalo, NY: State University of New York at Buffalo dissertation.
- Kwon, Song-Nim and Anne Zribi-Hertz. 2008. "Differential Function Marking, Case and Information Structure: Evidence from Korean," *Language* 84(2): 258-299.
- Lambrecht, Knud. 1994. *Information Structure and Sentence Form: Topic, Focus and the Mental Representations of Discourse Referents*. Cambridge: Cambridge University Press.
- Lee, Duck-Young. 2002. "The Function of Zero Particle with Special Reference to Spoken Japanese," *Journal of Pragmatics* 34(6): 645-682.
- Lee, EunHee. 2019. *Korean Syntax and Semantics*. Cambridge: Cambridge University Press.
- Lee, Hanjung. 2010. "Explaining Variation in Korean Case Ellipsis: Economy versus Iconicity," *Journal of East Asian Linguistics* 19(4): 291-318.
- Lee, Hanjung. 2015. "Case Particle Ellipsis," in Lucien Brown and Jaehoon Yoon, eds., *The Handbook of Korean Linguistics*, 196-211. Chichester: Wiley Blackwell.
- Lee, Hanjung. 2016. "Usage Probability and Subject-Object Asymmetries in Korean Case Ellipsis: Experiments with Subject Case Ellipsis," *Journal of Linguistics* 52(1): 70-110.
- Lee, Hanjung. 2017. "Effects of Information Status, Subject Type and Tense on Subject Case Ellipsis in Korean: An Experimental Study," *Discourse and Cognition* 24(4): 115-141.
- Lee, Sungbom. 2006. "A Pragmatic Analysis of Accusative Case-marker Deletion," *Discourse and Cognition* 13(3): 69-89.
- Lestrade, Sander and Helen de Hoop. 2016. "On Case and Tense: The Role of Grounding in Differential Subject Marking," *The Linguistic Review* 33(3): 397-410.
- Levshina, Natalia. 2021. "Communicative Efficiency and Differential Case Marking: A Reverse-engineering Approach," *Linguistics Vanguard* 7(s3): 20190087.
- Malchukov, Andrej L. and Helen de Hoop. 2011. "Tense, Aspect, Mood Based Differential Case Marking," *Lingua* 121(1): 35-47.
- Masunaga, Kiyoko. 1988. "Case Deletion and Discourse Context," in William Poser, ed., *Papers from International Workshop on Japanese Syntax*, 145-156. Stanford: CSLI Publications.
- Minashima, Hiroshi. 2001. "On the Deletion of Accusative Case Markers in Japanese," *Studia Linguistica* 55(2): 175-190.

- Park, Myung-Kwan. 2020. "Nominative Case Unmarking in Korean," *Studies in Modern Grammar* 108: 1-17.
- Park, Myung-Kwan and Arum Kang. 2020. "Korean vs. Chinese on Anaphoric Definites without a Determiner," Ms., Dongguk University and Hankyung National University.
- Rosch, Eleanor and Carolyn B. Mervis. 1975. "Family Resemblances: Studies in the Internal Structure of Categories," *Cognitive Psychology* 7(4): 573-605.
- Schwarz, Florian. 2009. *Two Types of Definites in Natural Language*. Amherst, MA: University of Massachusetts, Amherst dissertation.
- Shimojo, Mitsuaki. 2005. *Argument Encoding in Japanese Conversation*. New York: Palgrave Macmillan.
- Tsutsui, Michio. 1984. *Particle Ellipsis in Japanese*. Urbana-Champaign, IL: University of Illinois at Urbana-Champaign dissertation.
- Watanabe, Yasuko. 1986. "Two Kinds of Ellipsis in Japanese Discourse: A Quantitative Text Study," *Studies in Language* 10(2): 337-351.

Hanjung Lee, Professor
 Department of English Language and Literature, Sungkyunkwan University
 25-2 Sungkyunkwan-ro, Jongro-gu, Seoul 03063, Republic of Korea
 E-mail: hanjung@skku.edu

Received: November 16, 2021

Revised: December 9, 2021

Accepted: December 29, 2021