Editorial

The editors are pleased to welcome you to the tenth volume of FULL, an open access international journal providing a platform for linguistic research on modern and older Finno-Ugric and other Uralic languages and dialects. Since its inception in 2012 FULL has published comparative research as well as research on single languages, including comparison of just Uralic languages or comparison across family lines, welcoming both formal linguistic and empirically oriented accounts.

The first article in this volume by Katri Priiki, titled *From pronoun to particle: Finnish* tuo *'that' and* tuota *'well, erm'*, investigates the hesitation particle *tuota*, the partitive form of the demonstrative *tuo*. The question is why this demonstrative in particular has acquired the function of hesitation particle, and what the role of partitive case is in this process. The paper shows, based on data from a corpus of everyday conversations, that while non-referential discourse particle uses of the pronoun are also attested in the nominative case form, the partitive form is more frequent in this function. In this use the pronoun implies that the referent is only just becoming the target of attention, and partitive case is used with referents that are not fully individuated. It is proposed that the pragmaticizing of *tuo* may be a part of a more general phenomenon in which Finnish pronouns tend to turn to particles, particularly in their partitive form, and begin occurring at the beginning of a speaking turn.

The second article, by Asta Laugalienė, titled *Object marking with discrete objects in Finnish and Lithuanian*, investigates the case of direct objects, an issue that is notoriously complex in many Uralic languages, including Finnish. Lithuanian, a Baltic language, also exhibits an alternation between accusative and the partitive genitive, which, on the face of it, is similar to the alternation between accusative and partitive in Finnish. Comparing the semantic factors that give rise to the variation in the case-marking of discrete objects in Finnish and Lithuanian, Laugalienė finds that the Lithuanian partitive has some of the same NP-related functions as the Finnish partitive, but not the aspect-related functions. In Finnish the most important factor in object case licensing is the culmination of the event (or the presence/absence of the endpoint), while in Lithuanian quantification plays the most important role. This functional difference is related in the paper to the fact that Lithuanian, but not Finnish, has possibilities to express aspectual distinctions (like irresultativity) using verbal particles.

The last contribution is a methodological note by Pauli Brattico, titled Computation and the justification of grammars. In this piece the author revisits the original criteria proposed by Chomsky in Aspects for the justification of grammars and suggests that current computational methods provide a desirable and practicable tool for such purposes. It is argued that such computationally rigorous methods can help assessing observational, descriptive, explanatory and psycholinguistic adequacy of formal linguistic theories. The author supports the feasibility of this general approach to the justification of grammars by presenting a case study in the form of a Python model of Finnish agreement.

The current one is the final volume of our online journal. As of 2022, FULL will be incorporated in *Journal of Uralic Linguistics* (JUL), a new journal to be published by John Benjamins, with two of the editors of FULL functioning as editors-in-chief. While we regret to let go of FULL as an independent journal, we don't think it will be any

disadvantage either for the readers or for the many scholars who have published their papers in FULL. The ten published volumes will continue to be available on the journal's website as before, and there will also be a link to FULL on the website of JUL, which will remain there indefinitely. See https://benjamins.com/catalog/jul. JUL, very much like FULL, aspires to serve an integrative role in Uralic linguistics by striving to bridge the gaps between various research traditions and areas of specialization, providing them with a common platform. As a journal published by a leading international publishing house in linguistics, it aims to bolster the impact that results from the study of Uralic languages have on general linguistic theory and typology.

We take this opportunity to thank all colleagues who have contributed their articles to FULL over the last ten years, sharing their research with our readers. We are also indebted to all members of the editorial board and other colleagues who have generously devoted their precious time and expertise to reviewing for FULL.

Linguistic research on Uralic languages has been undergoing profound and multi-layered renewal and expansion in recent years. This shift has been marked by the extension of in-depth linguistic work on general linguistic topics of current interest to an ever-growing number of Uralic languages, as well as the appearance of electronic research tools. We look forward to the further development of this exciting and promising field in the years to come.

The Editors

From Pronoun to Particle: Finnish *tuo* 'that' and *tuota* 'well, erm'

Katri Priiki

This article studies the continuum of referential, vaguely referential, and particle-like occurrences of the Finnish demonstrative pronoun *tuo* 'that'. *Tuo* is peculiar among hesitator demonstratives, since it has pragmaticized to its partitive form *tuota* instead of its nominative form and it is not the same pronoun that is used in the function of a definite article (*se*). The article aims to shed light on the question of why this form in particular has pragmaticized to a hesitation particle. The results reveal that it is not only the partitive forms but also other case forms of the pronoun that may be used without a clear referent. The meaning features of the pronoun *tuo* imply that the referent is only just becoming the target of attention, and the partitive case is used with referents that are not fully individuated. When an abstract entity is referred to in partitive object role, the referentiality of a determiner or a placeholder may become unclear.

Keywords: *demonstrative pronoun, hesitation, discourse particle, pragmaticizing process, word search, Finnish*

1 Introduction

The sounds and words used to express hesitation and planning are part of a sporadic group that has been neglected in studies. Early studies on Finnish planning particles (Vuorinen 1981; Ravila 1945; Penttilä 1963) described them as semantically empty and without a syntactic function, often also stating that they should be avoided. However, in everyday conversation, planning particles are a very frequent phenomenon in all languages and they deserve to be thoroughly studied. In interactional linguistics, even the smallest parts of a language are considered to have an important function, even though their meaning may not be easy to describe.

This article focuses on the Finnish planning expression *tuota* 'well, erm', which originates from the demonstrative pronoun *tuo* 'that'. The partitive form of the pronoun, *tuota*, is frequently used non-referentially; in the latest descriptive grammar (Hakulinen et al. 2004: § 861), it is listed among discourse particles. It also often forms particle chains with other particles—for example, the chain *tuota noin* 'well, erm'—with the instructive form of the corresponding plural pronoun *nuo* 'those' (Etelämäki & Jaakola 2009: 191).

Earlier research on the particle *tuota* suggests that its functions include expressing hesitation, word search, the incompleteness of a turn, and self-repair (Penttilä 1963: 545; Lappalainen 2004: 128–131). The only thorough study focusing on the particle (Etelämäki & Jaakola 2009) reveals that *tuota* is not a means for the current speaker to reserve the turn for themselves but a genuine negotiation regarding who is going to speak next. They argue that the main semantic feature of *tuota* 'well, erm' is openness, which can be linked to both the meaning of the demonstrative pronoun *tuo* 'that' and to the meaning of the partitive case.

In this article, the focus is on borderline cases in which an occurrence of the words *tuo* or *tuota* can be interpreted as having either a function of a referential pronoun or a non-referential particle. By presenting a continuum from clearly referential use via vague referentiality to non-referential use, I show that the referentiality of almost any form based

on the demonstrative may not be very precise. I propose that the pragmaticizing of *tuo* may be a part of a more general phenomenon in which Finnish pronouns tend to turn to particles, particularly in their partitive form, and begin occurring at the beginning of a speaking turn. In this article, I focus on syntactic and semantic analysis, touching only lightly upon the prosody of the expressions. My preliminary observation regarding the prosody of *tuota* is that there is no clear pattern which would enable a differentiation between its referential and non-referential occurrences, and the phenomenon requires a thorough study.

I also link the use of the Finnish demonstrative pronoun *tuo* and the particle *tuota* 'well, erm' to earlier research on atypical uses of demonstrative pronouns in other languages. It is not a typologically uncommon phenomenon that a demonstrative pronoun is used as a filler word in spoken language when a speaker encounters trouble formulating his or her utterance. Demonstratives are shown to have this function in Japanese, Korean, Mandarin, Lao, Romani, Russian, and Spanish (Hayashi & Yoon 2006). However, this function is often forgotten when demonstratives are discussed. Both Hayashi and Yoon (2006) and Etelämäki and Jaakola (2009) suggest that, even in highly pragmaticized functions, demonstratives continue to retain a certain degree of indexicality. Hayashi and Yoon (2006) also argue that the features that make demonstratives, among all linguistic devices, suitable for expressing hesitation are their pointing function and the aspects of participant access that they express. While the other two Finnish pronouns that, in partitive form, have a particle function (*sitä, häntä*) have their own meanings linked to those of the corresponding pronouns, *tuo* has suitable semantic features to be considered a hesitation particle.

The data of the study comes from Arkisyn, the morphosyntactically annotated corpus of everyday Finnish conversations. I examine conversations from the viewpoints of interactional linguistics and emergent grammar, following Ford's (1993) thoughts that grammar emerges through interaction among participants who are constantly reusing and modifying prior utterances to achieve current interactive goals.

When referring to the reanalysis of the pronoun to a particle (e.g. in Section 2.2), I use the term *pragmaticization* instead of *grammaticalization* to express that the process does not involve the emergence of new grammatical markers (for discussion on these two terms, see e.g. Heine 2013: 1217–1120).[1] *Pragmaticization* (or *pragmaticalization*) has been defined as a process by which a unit changes its propositional meaning in favour of an essentially discourse interactional meaning (Frank-Job 2006: 397; Hayashi & Yoon 2006). While grammaticalization tends to lead to syntactic integration, pragmaticization involves, for example, increased syntactic freedom, semantic-pragmatic scope, and optionality (Heine 2013: 1218). These are central features of the Finnish particle *tuota* 'well, erm' also.

In the next section, I describe the Finnish demonstrative system, earlier research on the particle *tuota*, and the studies on demonstratives that are used to express hesitation in other languages. In Section 3, I present my data. Then, I introduce the use of *tuo* and *tuota* as placeholder demonstratives and hesitation particles in the data (Section 4) before concentrating on the borderline cases with unclear referentiality (Section 5). In Section 6, I discuss the relation of *tuota* to other Finnish particles originating from partitive forms of demonstratives, and in Section 7, I conclude the findings and their implications.

[1] Heine (2013) suggests the term *cooptation* to describe the process.

2 Background

The Finnish language has three demonstrative pronouns, *tämä* 'this', *tuo* 'that', and *se* 'it; 3SG'. The same forms are used as determiners of a noun. All these pronouns also have plural variants (*nämä* 'these', *nuo* 'those', *ne* 'they') and they inflect in 12 cases.[2] The most frequent case forms of *tuo* in singular and plural and their standard forms and colloquial variants, which the examples of this article mostly represent, are presented in Table 1. Added to these forms, numerous kinds of demonstrative adverbs and proadjectives (pronominal forms used like adjectives) are derived from them.

	Singular	**Plural**
Nominative	*tuo (toi, tua)*	*nuo (noi)*
Genitive	*tuon (ton, tuan)*	*noiden, noitten*
Partitive	*tuota (tota, tuata)*	*noita*

Table 1: *The standard and colloquial forms of the Finnish pronoun* tuo *in the most frequent cases*

In 2.1, I present earlier studies on how the pronoun *tuo* differs in meaning and in use from the other two demonstrative pronouns. In 2.2, I discuss the earlier observations on the connection of the pronoun *tuo* and the particle *tuota* 'well, erm' that is pragmaticized from the partitive form of the pronoun. In 2.3, the last subsection, I present the continuum from placeholder demonstratives to fully pragmaticized hesitation particles—that is, the typological context where I place the Finnish *tuo* and *tuota* in this article.

2.1 *Tuo* 'that'

Tämä 'this' is traditionally considered proximal to the speaker and *tuo* 'that' as distal or proximal to the hearer (Larjavaara 1990). A recent experimental study (Reile et al. 2019) reveals that with physical objects as referents, speakers refer to targets that are further away significantly more frequently with *tuo* compared to *tämä*. However, examining conversational data has shown that when there is no apparent spatial contrast, the frequency of usage of *tuo* compared to *tämä* is instead explained by cognitive, social, and affective factors (e.g. Östman 1995; Laury 1997). In contrast, *se* 'it; 3SG' is a neutral anaphoric pronoun. In informal speech, it is the most common way of referring to any kind of referent, even people, although standard Finnish has a separate third-person personal pronoun *hän* 'he, she'. In numerous languages, pronouns characterize referents as, for example, animate or inanimate, male or female. However, the Finnish demonstrative pronouns only imply that the entity referred to is a thing or a person, rather than a quality, location, manner, or time. The latter may be referred to with proadjectives (e.g. *tuollainen* 'that kind') and pronoun-rooted adverbs (e.g. *tuolla* 'over there', *noin* 'that way', *tuolloin* 'at that time').

According to Etelämäki (2006, 2009), the main semantic features of *tuo* are referential openness and indexical unmarkedness. Referential openness implies that when *tuo* is used, the process of identifying the referent is still ongoing. The referent may, for example, become the target of attention the moment the reference is uttered, not before. This contrasts with another Finnish demonstrative pronoun, *se*, which implies that the referent

[2] Demonstrative pronouns are not used in the abessive or comitative cases.

is already known by all participants of the conversation. In contrast, indexical unmarkedness implies that the referent is equally (non)accessible for both the speaker and hearer (e.g. both having or not having a visual contact or memory recollection). This feature separates *tuo* from the third demonstrative pronoun, *tämä* 'this', which indicates that the speaker has the primary access (Etelämäki 2006, 2009).

Tuo points rather than presents—that is, the reference directs attention to the referent but the host utterance does not give it a new interpretation, as it does when *tämä* is used (Etelämäki 2009). In certain contexts, *tuo* also expresses a contrast or a figurative distance between its referent and another subject or the speaker (Laury 1997; Priiki 2015). Similarly to other Finnish demonstratives, it is occasionally used in the tail (or right-dislocation) construction, as in example (1) (Priiki 2020).[3] In the Finnish tail construction, the same referent is first referred to with a demonstrative pronoun and then a second time with a full noun phrase that usually has the same demonstrative as a determiner. In the example below, the first pronoun, the placeholder, is bolded and the full noun phrase, the tail, is underlined.

(1) ***Toi*** *hoitaa* *sitä* *toi Martta* *nyt.* (Priiki 2020: 182)
that takes.care 3SG.PART that NAME now
'That (woman) takes care of her now, Martta.'

In the tail construction, *tuo* is used as a placeholder demonstrative, a function that I will discuss in Section 2.3, below. Apart from other functions (cf. Priiki 2020), a tail construction may be used to postpone the lexical reference in a situation in which the speaker has trouble finding the appropriate definition. In Finnish, any argument that may be the topic of the sentence—that is, not only a subject—can be the target of the double-reference in this kind of structure (Holmberg & Nikanne 2002: 71). Even though all three demonstratives occur in tail constructions, particularly the variant with the pronoun *tuo* is used in contexts with markers of word search and hesitation. I suggest that analysing this kind of use further may help to explain why a form of *tuo* has pragmaticized to a hesitation particle.

2.2 *Tuota* 'well, erm'

In her sociolinguistic study, Lappalainen (2004: 118) notes that there are significant personal differences in the number of hesitation particles used in general, and people tend to prefer one or another variant. She speculates that hesitation particles in general—and particularly *tuota* 'well, erm'—are currently spoken more frequently than they were a few decades ago (Lappalainen 2004: 113). However, this is difficult to prove due to the lack of comparable conversational data. In any case, the particle use of *tuota* is not a new phenomenon as such: linguists already made notes about it in the nineteenth century (Lönnbohm 1879; Setälä 1883; Latvala 1894). In dialect interview data recorded in the

[3] The glosses used in the examples: 2SG = second person singular, 3SG = third person singular, ADE = adessive, CLI = clitic, ELA = elative, GEN = genitive, ILL = illative, INE = inessive, MS = misspelled item, NEG = negation verb, PART = partitive, PCP = participle, PL = plural, PTC = particle, TRANSL = translative. In the transcripts, a question mark expresses rising intonation and a period indicates falling intonation at the end of an utterance. In the translations of the examples, the use of ending marks follows the norms of written language.

1960s, there are occurrences of the particle *tuota* in all regional dialects, with an emphasis on eastern dialects (Pihlaja 1971).

Lappalainen (2004: 128) also speculates that the context in which the pragmaticization of the partitive form of the demonstrative pronoun was initiated might be its use as a determiner of a noun (e.g. *tuo tyttö* 'that girl'). I discuss this claim later in this article. A word pragmaticizing to a particle is described with reductions of phonology, morphology, and syntactic behaviour; in addition, a prototypical particle is non-compositional and short and has a non-restrictive and rather procedural meaning (Hakulinen & Seppänen 1992: 535–537; Heine 2013: 1209). Many of these criteria are listed as central features of particles in the latest comprehensive Finnish grammar (Hakulinen et al. 2004: § 794), which also mentions that particles cannot be targets for a question, negation, or focus, are never inflected, and cannot have or be determiners. The data used in this article includes numerous occurrences of *tuota* 'well, erm' that do not fulfil all these criteria. In principle, *tuota* is an inflected form and it may occur in such syntactic positions that it could be interpreted as a placeholder or a determiner of a noun, even though the exact referent may be unclear. In addition, the prosodic behaviour of *tuota* does not form a clear pattern. Even though many occurrences are either prosodically set off from the rest of the utterance, or they have reduced prosodic prominence or, for example, a lengthened final syllable, the data also includes non-referential occurrences with very clear stress (e.g. turn-initial *tota noin* chains). On the other hand, the prosody of clearly referential occurrences may resemble particles in word search contexts.

In her study, Lappalainen (2004: 114) has counted only unambiguous occurrences of the particle *tuota* 'well, erm'. She observes that, in her data, interpretation as a particle is usually clear because the word occurs in a position where a pronoun would not be possible or the word does not inflect in case or number like a pronoun would. In (2a–c), I present simple examples of a clearly non-referential particle, a typical pronoun, and a vague case falling between both these categories, respectively.

(2) a. *No saa=ks* ***tuo-ta*** *os-i-ks.* (D131)
well can-CLI that.PART part-PL-TRANSL
'Well can you put that one to parts?'

b. *Se autto meit jossain noissa mm* ***tota*** *editoinne-i-s.*
3SG helped we.PART some.INE those.INE PTC ehm editing-PL-INE
'S/he helped us in some- those- ehm editings.' (SG124)

c. *Meiä* ***tota*** *katuu ei ollu aura-ttu vielä?* (SG151)
our that.PART/ehm street.PART NEG be.PCP clear-PCP yet
'Our- that/ehm street had not been cleared (of snow) yet.'

A clearly referential occurrence of the pronoun *tuo* in partitive is presented in (2a). The pronoun functions as the object of the clause and refers to a concrete object (a puzzle toy) that lies on the table in front of the participants. The speaker touches the object during his utterance, and pronoun *tuo* contrasts the referent with another similar object they discuss. In contrast, the second example (2b) presents a context where the interpretation of *tuota* as a particle is the only possible option. The word *tuota* occurs in the middle of a noun phrase, between a determiner demonstrative and a noun. The noun and its determiner are inflected in plural inessive—the singular partitive form *tuota* cannot refer to the same target. Actually, the determiner used here is the plural form of *tuo* and choosing it for the determiner is another way of expressing hesitation in (2b). There is also a hesitation in the

sound (*mm*). In this example, the function of *tuota* is to delay the production of the word *editoinneissa* 'in editings', which the speaker is searching for. The third example (2c) represents a case in between these two ends, the type that is the focus of this article. In (2c), the context of *tuota* resembles the one in (2b), but the word following it is inflected in the same case and number—singular partitive—and *tuota* can be interpreted as the determiner of the noun *katuu* 'street', produced as a self-repair to substitute the genitive determiner *meiä* 'our'. However, the particle interpretation is supported by the fact that *tuota* is unstressed and uttered very fast; moreover, in the context of (2c), it appears to have a rather imprecise meaning that rather projects a story that is to begin rather than being a determiner for the noun 'street'.

Etelämäki and Jaakola (2009) investigated the meaning of the particle from the viewpoint of cognitive linguistics. Their study, based on a relatively small data set, proposes a schematic meaning for the particle: they argue that *tuota* 'well, erm' expresses openness on various levels; in other words, the participants are negotiating something. This could be, for example, the form of a reference to a certain target or the entire speech act that is about to follow. Etelämäki and Jaakola (2009) do not comment on the development of the particle other than indicating the link among the meanings of the particle, the pronoun, and the partitive case. They ignore the possible nominative forms among the non-referential uses, do not discuss ambiguous cases, and only examine occurrences where *tuota* 'well, erm' is not accompanied by other particles.

The Finnish partitive ending *-(t)a/-(t)ä* used to be a separative case, indicating movement away from something. Since it lost its locative meaning, it transformed into a partitive. The Finnish partitive case has multiple uses; among other things, it expresses quantification and aspectual distinctions, particularly unboundedness (Huumo 2010; Larjavaara 2019). Huumo (2010: 95) mentions that the general function of the partitive is to indicate incompleteness. In contrast, according to Helasvuo (1996: 13), what connects the different uses of partitive in interaction is low transitivity, when transitivity is understood as a feature of the entire clause rather than an individual verb. In conversational data, partitive noun phrases (NPs) are often mass nouns or refer to inanimate targets. Thus, they are less individuated than are objects in the accusative or nominative cases. Referents that are introduced to a conversation with an NP in the partitive case are not usually mentioned again and they are not central to communication (Helasvuo 1996: 28–30).

Tuo is not the only Finnish pronoun that has pragmaticized to a particle in singular partitive form. The partitive form of another demonstrative *se* (*sitä*) is frequently used in particle function; moreover, the third person pronoun *hän* (*häntä*) has similar, even though less frequent, uses in spoken dialects. I discuss the relationship among these three particles and the partitive case in greater detail in Section 6.

2.3 Placeholder and hesitation demonstratives

In this subsection, I compare the Finnish *tuo* with demonstratives used to express hesitation in other languages. In the context of word-formulation difficulty, there are three distinct usage types of demonstratives, described by Hayashi and Yoon (2006): the placeholder use, the avoidance use, and the interjective hesitator use. The placeholder use and the interjective hesitator use are relevant where the Finnish pronoun *tuo* is concerned.

Placeholders are referential and participate in the syntactic structure of the utterance—that is, those forms of demonstrative pronouns are used that correspond syntactically and semantically to the word for which the pronoun is functioning as the

placeholder. In contrast, a pronominal form used as an interjective hesitator is not referential, has no role as a clausal constituent, and usually has little correspondence to the word a speaker is searching for. The function of a placeholder is to advance the progress of a syntactic structure that is being produced by filling a required slot. The lexical reference is then produced later, often as an independent nominal phrase, which connects the placeholder use to self-repairs and tail constructions. In contrast, an interjective hesitator merely delays the production of the remainder of the utterance, thereby signalling that the speaker aims to continue their turn.

Placeholder use may resemble the cataphoric use of demonstratives, where a pronoun refers forward to a lexical noun phrase that is about to follow. However, using placeholders is motivated by constraints in cognitive processes, such as difficulty in remembering a word when it requires articulation. Thus, the use of placeholders is different from the cataphoric uses of demonstratives in terms of motivation. In certain languages, separate pronouns are used in these functions: in Japanese and Korean, proximal forms are used in cataphora, distal, and medial forms as placeholders (Hayashi & Yoon 2006). If the same variants are used in both functions, like in Finnish, cataphoric references may result in similar structures as using placeholders: a first mention with a demonstrative pronoun is subsequently followed by a lexical noun phrase. The difference is that cataphoric references are planned structures. In Finnish, for example, tail constructions may be used to modify the word order and information structure of an utterance by presenting a long lexical phrase at the end of the utterance, where new information is usually presented (Priiki 2020). Of course, it is not possible to know the motivation for a certain linguistic structure for sure. However, if a first-mention demonstrative is accompanied by markers of hesitation, such as pauses and hesitation sounds, we may assume that its use is at least partially motivated by difficulties in lexical retrieval.

Whether proximal or distal demonstratives are used as placeholders varies in different languages. In Japanese, the forms used are the distal variants in the three-part distance-based system. In Korean, distal and medial forms may be used. In Mandarin and Indonesian, with two distance categories for demonstratives, both distal and proximal demonstratives are used as placeholders. Moreover, the kind of an entity a placeholder can project varies. In Indonesian, a placeholder demonstrative may substitute linguistic items on various levels: it may even be used instead of a verb root. Occasionally, placeholder demonstratives form fixed expressions with certain other words. For example, in Mandarin, the distal placeholder *na-ge* is often followed by the word *shenme* 'what' (Hayashi & Yoon 2006). Finnish *tuo* also, when used in hesitation and word search, often occurs together with certain adverbs and particles—for example, *tuota noin* 'well erm'.

When interjective hesitators are pragmaticized for the function, they diverge from ordinary demonstratives for syntactic distribution, referentiality, and correspondence between morphology and semantics. In other words, they turn into discourse particles, which have more distributional freedom than the original demonstratives. Unlike referential placeholders, interjective hesitators can appear anywhere during an utterance-in-progress. For example, Japanese *ano* is an adnominal demonstrative and must be placed before a noun, but as a hesitator it can appear anywhere (Hayashi & Yoon 2006: 507). Similarly, Finnish *tuota*, a singular partitive form, may appear as an interjective hesitator in contexts where the word searched for is in plural form and inflects in some other case.

Interjective hesitators may acquire functions that pragmaticize further from their use in word searches. In Japanese and Korean, hesitator demonstratives often preface the introduction of a new topic or an initial action, like a proposition (Hayashi & Yoon 2006:

528). In Russian, too, the compound hesitator *eto samoe* 'this very' is often used at the beginning of a turn (Podlesskaya 2010: 20). In this kind of context, the function of interjective hesitators is to draw the hearer's attention to the next action and to give a hint of how to interpret the utterance. Moreover, Finnish *tuota* is said to have other functions than merely word search. According to Etelämäki and Jaakola (2009), when it occurs at the end of a turn or forms a whole turn alone, it indicates that the next action is not yet decided.

Both Hayashi and Yoon (2006) and Etelämäki and Jaakola (2009) suggest that, even in these highly pragmaticized functions, the interjective hesitator demonstratives retain a certain degree of indexicality. Hayashi and Yoon (2006) also argue that the features that make demonstratives—among all linguistic devices—suitable for expressing hesitation, are their pointing function and the aspects of participant access that they express. For example, in Korean, different demonstratives used in word searches invite a different kind of participation: forms that propose shared access to the referent, invite emphatic reactions, or collaborative word search. In turn, when speaker-centred forms are used, the recipient is passive (Hayashi & Yoon 2006: 516–517). As mentioned above, the Finnish pronoun *tuo* implies that the referent is not currently the centre of attention but is accessible independently by all participants, not just the speaker.

It is evident from a variety of languages that the same forms are often used as hesitation particles and definite articles, thereby invoking the sense of 'you know what I'm talking about'. This is the case in Estonian, a language closely related to Finnish, where the pronoun *see* may express word search (Keevallik 2010). However, in Finnish, the definite article, used only in colloquial language, is the anaphoric demonstrative *se* (Laury 1997). *Tämä* 'this' and *se* 'it; 3SG' may also occasionally occur as placeholders in word searches, but only *tuo* 'that' has extensive, conventionalized, and pragmaticized use as a hesitator demonstrative. Further, in Swedish spoken in Finland, both *dedär* 'that' and *dehär* 'this' have conventionalised to word searches (Wide 2011). Their pragmaticization is said to have Finnish influence, even though in Finnish, only *tuo,* which corresponds to *dedär* (and not to *dehär* that translated as *tämä*), is frequently used in word searches.

In Finnish, the partitive form, in particular, has pragmaticized to a hesitator. This is noteworthy because usually the default form that projects a referent in some other number and case is the singular nominative form. For example, in Russian, the default hesitator demonstrative is *eto*, which is a nominative singular neuter form (Podlesskaya 2010). However, Podlesskaya (2010) notes that other demonstrative forms are used as placeholders when they correspond to the word searched. She also provides an example where a placeholder is in accusative without a corresponding referent, thereby interpreting this to be due to an elliptic verb (ibid.: 21); moreover, she provides an example from nineteenth century Russian, in which it was possible to use a genitive form of a distal demonstrative (*togo*) as this kind of hesitator demonstrative (ibid.: 19). In Section 5.2, I reflect on the possibility of elliptic structures also being responsible for the partitive form of the Finnish *tuota*.

3 *Tuo* and *tuota* in the data

The data for the study is the morphosyntactically annotated corpus of conversational Finnish, Arkisyn. It comprises approximately 40 hours of naturally occurring video- or audio-recorded and transcribed conversations. In the corpus, there are 4,066 occurrences of the pronoun *tuo* or the particle *tuota* 'well, erm'. Of these, 907 are different variants of

the nominative form (*tuo, toi, tua*), while 1,646 are variants of the partitive form (*tuota, tota, tuata*). The remainder are forms inflected in other cases. Further, 2,553 of the occurrences are coded (by the coders of the corpus) as different forms of the demonstrative pronoun, while 1,513 are interpreted as particles.

However, with regard to *tuo* and *tuota*, the coding of the corpus is unstable, which reflects the fact that, in this case, the line between a pronoun and a particle is not straightforward. The fact that only the partitive form is listed in grammar as a particle has affected the coding, thereby causing the coders to likely interpret the partitive forms as particles and other forms as pronouns; however, this solution is not entirely systematic. In this article, I do not suggest that the nominative form or any other case forms apart from the partitive must be considered to belong to the class of particles. Instead, my aim is to show that the other forms can be used non-referentially in the particle function, as in the next example. In (3), three girls are doing their homework together. One of them repeats the nominative form *toi* multiple times. During her turn in (3), the speaker is fidgeting around, not able to concentrate, and the words she produces are mostly nonsense. Apparently, she is looking for some new topic for conversation. The context, or checking the video recording, does not provide any clues for a possible referent.

(3) SG120[4]

01	*Hmm::?*	*(minä*	*olen*	*niin)?*	*hm::*	*hm::;*	*(0.5)*	*°no*	*voih,°* *(.)*
	PTC	I	am	so	PTC	PTC		well	oh
02	*.hh*	***toi***	***toi***	***toi***	***toi,***	*(.)*	*°(pum*	*pum*	*pum)°, (1.5)*
	BREATH	that	that	that	that		boom	boom	boom

Moreover, the *toi* forms here do not belong to any syntactic structure; they cannot be interpreted in as straightforward a manner as placeholder demonstratives. At the most, they could be holding place for some abstract topic of conversation that the speaker is searching for. Thus, their function is that of hesitation particles, even though the form is nominative.

The number of pronominal occurrences of *tuo* (only singular pronouns) in different case forms is presented in Table 2. Because demonstrative adverbs behave differently from demonstrative pronouns—for example, their inflection paradigm and the ability to function as determiners are limited—the numbers are separated from pronominal occurrences. There are 2,067 pronominal occurrences and, of these, 666 function as determiners of nouns. The nominative and partitive cases are a target of interest and together with the genitive case, they are the most frequent cases. The other cases are grouped together in Table 2.

The number of case forms rests on the annotation of the corpus, and the coding may contain other errors in addition to the ones reported above. While I have corrected mistakes when encountering them in data searches, I have not checked each of the occurrences personally. However, I believe the coding is sufficiently accurate to provide an overview of the distribution of different forms in the data. In Table 2, it is evident that the partitive case is not a very common case for the demonstrative pronoun *tuo*: it only

[4] The transcription symbols used in the following examples are: lengthening of a sound; ? rising intonation; . falling intonation; , levelled intonation; (.) a short pause, .hh inhale, (1.5) a pause longer than 0.5 seconds (length mentioned); ° whispered part; # creaky voice; £ smiling voice; @ altered voice; ↑ high pitch; ↓ low pitch; [overlap; >< fast tempo; <> slow tempo.

constitutes 6% of the occurrences. For nouns in Arkisyn, the share of the nominative case is 37%, partitive case 20%, and genitive case 11%. However, among *tuo* determiners, the share of the partitive case is closer to the nouns (11%). In Finnish, the determiners always inflect in the same case as the main word. In Section 5.2, I reflect upon whether the use of the determiner is a possible context for the pragmaticizing development of *tuota*, as Lappalainen (2004: 128) speculated.

Case	**All occurrences**	**Pronominal occurrences**	**Adverbs**	**Determiners**	**Independent pronoun phrases**
Nominative (*tuo*)	1,060 (42%)	1,060 (51%)	0	342 (51%)	718 (51%)
Genitive (*tuon*)	165 (6%)	165 (8%)	0	76 (11%)	89 (6%)
Partitive (*tuota*)	146 (6%)	146 (7%)	0	73 (11%)	73 (5%)
Other cases	1,182 (46%)	696 (34%)	486	175 (26%)	521 (37%)
All	2,553	2,067	486	666	1,401

Table 2. *The pronoun* tuo *in singular in different cases in Arkisyn*

The tendency in the coding of the corpus appears to be that vague cases are interpreted as pronouns rather than particles, not the other way around. In this article, these vague cases are the target of interest. From the automatic search results[5] of both lemmas, *tuo* 'that' and *tuota* 'well, erm', I have collected all such occurrences where the interpretation of the referentiality of the word is not clear by judging the transcribed context that is visible in the search results and listening the recording of the collected utterances. My collection includes 318 examples, which means that approximately 8% of the occurrences of *tuo* and *tuota* may be ambiguously referential. However, the interpretation of the referentiality is subjective at least to a certain extent and, added to this, the number of vague cases has significant variation among different recordings that could be related to topics that are discussed as well as to personal strategies of expressing hesitation. Providing reliable quantitative observations of the phenomenon would require a more extensive study. Thereafter, I have inspected the broader context and the original video recordings of examples that represent the types recurring in the data.

4 *Tuo* and *tuota* as placeholders and hesitators

In this section, I describe how the Finnish demonstrative pronoun *tuo* is used in the data as a placeholder and hesitator demonstrative—that is, the functions presented above. The examples below represent typical cases documented in earlier studies: (4) and (5) are cases where *tuo* functions as a demonstrative pronoun but expresses difficulties in retrieving a lexical reference (see Priiki 2015, 2020), while (6) represents the partitive form *tuota* as a

[5] Command [(lemma = "tuota") | (lemma = "tuo")] in Korp search interface.

fully pragmatized particle that has lost inflection and referentiality (see Etelämäki & Jaakola 2009). The vague cases, where the division into referential placeholder function and non-referential hesitator function is not clear, are discussed in Section 5.

The demonstrative pronoun *tuo* may be used as a placeholder in, for example, the tail construction, presented above in (1). Other demonstrative pronouns are also possible and frequent in tail construction placeholders; however, in the Arkisyn corpus, among tail constructions with *tuo*, in particular, there are occurrences where a speaker has trouble identifying the appropriate lexical definition and, occasionally, the lexical description of the referent stays missing. Priiki (2020) focuses on cases where a tail construction with a *tuo* placeholder is used as a means of modifying the word order and information structure. This article continues the study by focusing on such occurrences where hesitation and processing trouble appear to be a plausible explanation for choosing a *tuo* placeholder. The majority of the *tuo* placeholder demonstratives in the Arkisyn data are in the nominative case, but other case forms also occur.

In (4), two women, Iina and Ritva, are discussing a new curtain fabric Iina has bought. At the beginning of the excerpt, Ritva refers to the fabric with the pronouns *se* 'it; 3SG' and *tuo* 'that'. *Tonne välii* 'into there between' refers to a ventilation gap.

(4) SG446

01 Ritva: *Ni et se on sen verran pidempi*
so that 3SG is it.GEN much longer
(1.3)

02 *että se ei jää* ***tonne*** *#välii sitte* ***toi****#,*
that it NEG stay over.there between then that
'So that it (the fabric) is that much longer so it won't go there in between, that,'

03 Iina: *Joo. (.) .Hhh (1.7) no mut* ***toi****=han on tosi syvällä* ***toi*** *(.)*
yeah BREATH well but that=CLI is really deep.ADE that

04 *siis* ***toi*** *t- ikkuna et*
I.mean that window that
((shows a measure with hands))

05 *ei=hän se tuu sinne se verho.*
NEG=CLI 3SG come there the curtain
'Well but that is so deep that, I mean that t- window so it won't go there, the curtain.'

When Iina utters *toihan*, which is the nominative form with a clitic particle *-han*, referring to the window, a pronoun reference alone is difficult to interpret because there are several possible referents (the window, the fabric, the ventilation gap). Iina adds another *toi* and after a slight pause yet another; then, she begins a word with a *t*-sound. However, the noun she finally produces, *ikkuna* 'window', does not begin with a *t*-sound. The formulation trouble here is probably caused by difficulty in briefly describing a situation where, due to, for example, a thick wall a curtain hangs rather far from a window. 'Deep' is not an adjective that is usually used to describe windows. Iina uses placeholders to acquire more time to decide which word to use and to signal to Ritva that the choice of description may be somehow problematic.

Based on the video recording, both Iina and Ritva appear to be situated rather far from the window and even though Ritva makes a few gestures towards the window during

the utterances of lines 1 and 2, these references are not accompanied by exact pointing. In turn, Iina is holding the fabric in one hand during her utterance; producing the first *toi*, she looks at the window and while uttering the latter part of her turn (from *toi t- ikkuna* onwards), she gestures and returns her gaze towards Ritva.

In (5), a speaker is in conversation with a friend on the telephone, discussing a student party she has attended. She appears to have trouble choosing how to explain the location of the party to the friend, who does not know the city where the speaker lives. She first uses *tossa* 'over there', an inessive form of *tuo*, and then replaces it with the name of the place, which has a demonstrative proadjective as a determiner (*semmosessa* 'in [something] like it'). After this, she uses another demonstrative proadjective (*sella[se]ssa* 'in [something] like it') and two more inessive forms of *tuo* (*tossa*, *tos* 'there' or 'in that'), the latter of which is the determiner of the noun *keskusta* 'centre'.

(5) SG113

01	*Et*	*ne*	*oli*	***tossa***	*semmosessa*	*Driimissä*	*sellassa, (.)*	*pupissa*
	so	they	were	that.INE	sort.of.INE	NAME.INE	like.it.INE	pub.INE
02	*(.)*	***tossa***	*mhh*	***tota***	*noin*	***tos***	*keskustassa*	*ja (.) .h*
		that.INE	PTC	that.PART	PTC	that.INE	centre.INE	and BREATH
03	*me*	*oltii*	*sittes*	*sillä*	*Ullalla*	*aluks*	*ja*	*hh*	*me mentii*
	we	were	then	the.ADE	NAME.ADE	first	and	BREATH	we went
04	*sitt↑es*	*sii↓tä*		*siihe*	*pupiij - -*				
	then	from.there		the.ILL	pub.ILL				

'So the party was [lit. 'they were'[6]] in a kind of a pub, Driimi, over there, well, there in the centre, and first we were at Ulla's place and from there we went to the pub - -'

Among the referential forms in (5), there also occurs the non-referential partitive form *tuota*, which is accompanied by another particle, *noin*. While the inessive forms convey a location (perhaps a pub or a district), the particle chain *tuota noin* only indicates that the utterance is going to continue. Typically, the particle *tuota* occurs in the middle or at the beginning of a turn and it is more often accompanied by other particles than alone (Etelämäki and Jaakola 2009: 191–193).

Example (5) includes numerous different markers of hesitation, vagueness, and processing trouble. The demonstrative proadjectives emphasize the type and features of the referent instead of precise identification and they occur in word searches (Hakulinen et al. 2004: § 1411). There are several short pauses and both referential and non-referential occurrences of the pronoun *tuo*. The next excerpt (6) provides another example of the use of the partitive form *tuota* as a hesitation particle. The element postponed may also be a whole turn or action, as in (6). The excerpt is from a telephone conversation, where two friends have just decided the time they will meet on the next day. There are two occurrences of *tuota* in (6), one produced by each speaker. Both of them are accompanied by hesitation sounds and the first one is accompanied by another particle *noin*. The latter occurrence is at the beginning of the turn, which is a typical place for a *tuota* particle.

[6] In Finnish, words expressing numerous kinds of events, including parties—for example, *juhlat*, *bileet*—occur in plural even though they refer to a singular event (Hakulinen et al. 2004: § 558).

(6) SG111
01 P: *Lähen kolmen junalla et se on puol viis siel [lä.*
'I will leave on three o'clock train it will be there at half-past four.'
02 E: *[Joo.*
'Yep.'
03 P: *.Hhh (.) jep. Hh* ***tota*** *noi. Tmhh no.*
BREATH PTC BREATH that.PART PTC PTC well
'Okay, erm well.'
04 P: *.Mt [hh*
PTC
05 E: *[.Mtghh* ***tota*** *mth (.) soitat=ko sit huomenna.*
PTC that.PART PTC call.2SG=CLI then tomorrow
'Will you then call tomorrow?'

Etelämäki and Jaakola (2009) suggest that the particle *tuota* occurs in contexts where it is undecided how the conversation is going to proceed and that the place of the particle affects what is interpreted as 'open'. In word searches, *tuota* precedes a certain phrase or constituent which is being searched for or formulated. At the beginning of a turn, *tuota* may project certain new activity: in (6), the speakers negotiate whether the call is already about to end and if they still have something to discuss.

In this section, I have compared the referential placeholder use and the particle use of *tuo* and its partitive form. Examples from the Arkisyn corpus presented here confirm the earlier observations made by Etelämäki and Jaakola (2009). These provide a starting point for the examination of such occurrences in the next section, where an ambiguous interpretation of either the function of a referential pronoun or that of a non-referential particle is possible.

5 Ambiguous uses

The focus of this article is on ambiguously referential or non-referential uses of the pronoun *tuo*. In this section, I highlight several phenomena that recur in the data often enough to attract attention. Analysing the ambiguous occurrences reveals that referential vagueness may be approached from at least two perspectives. First, there are occurrences where the word appears to project an entity like a placeholder does, but the target of the reference in the context is somehow fuzzy and ambiguous. I examine these kinds of occurrences of different forms of the demonstrative pronoun *tuo* in Section 5.1 in order to answer the question of what makes this pronoun in particular likely to express hesitation.

On the other hand, there are cases where it is unclear whether the function of a word is a pronoun or a particle because it occurs in a syntactic position that is possible to interpret as part of a structure—a determiner of a noun, an object, or a subject. However, certain other features, like pauses, create an impression of non-referentiality. I examine these phenomena in Section 5.2.

5.1 A specific but non-salient referent

In Section 4, I showed how the forms of the Finnish demonstrative pronoun *tuo* may be used in the functions of a referential placeholder and of a non-referential hesitation

particle. By using *tuo*, a speaker also implies that the referent is not salient in the ongoing action, for example, storytelling. If the time gained with the placeholder is not sufficient for accessing the lexical phrase, the referent may be left non-explicated without problems in interaction. However, *tuo* also implies that the referent is accessible if necessary and, occasionally, another participant in the conversation wishes to clarify the reference.

In the data, almost any form of the pronoun *tuo* may be used with a non-explicated referent, as (7) shows. The example is from a telephone conversation between a mother and daughter, who are talking about an eye problem the daughter's dog is having. In Irja's turn, there are three different forms of *tuo*—a singular elative form (*tosta* 'from that'), a plural adessive form (*noil* 'on those'), and a demonstrative-rooted locative adverb (*tuol* 'over there'). The reference is not explicated and possibly unclear in all three forms, but this causes no problems for the co-participant in the conversation.

(7) SG124

01 Irja:	>*Mä*<	*katon*	***tosta***	*kun*	*mul=han (.)*	***noil=ha***	*(0.7)*
	I	look.1SG	that.ELA	because	I.ADE=CLI	those.ADE=CLI	
02	*oli*	***tuo(l)***	*se*	*Koira-n*	*ensi-ap#u:#*	*se*	*Pelle Nah-*
	was	over.there	the	dog-GEN	first-aid	the	NAME
03	*Pelle*	*Akselsson-in.*					
	NAME	NAME-GEN					

'I will check that (one), because I- they had over there the First-aid for Dogs, that (one) from Pelle Akselsson's.'

04 Heta: *Niih.*
'Yeah.'

05 Irja: *Niist luennoista. Mä katon nyt vielä sitten mitä vois silmissä olla muuta.*
'Those lectures. I will now check what else could be in eyes.'

In (7), the first non-explicated referent is a book, a leaflet, or maybe just a pile of lecture notes from a course. The reason for the unclear reference may be that the speaker cannot decide which description to choose. Thus, the use described here is similar to the use of the pronoun in word searches as a placeholder demonstrative (cf. ex. (1)). The same entity is subsequently referred to with the term *Koiran ensiapu* 'First-aid for dogs', marked known to the hearer with the article-like demonstrative determiner *se* (regarding the definite article in Finnish, see Laury 1997). Moreover, the past tense in the clause ('I-they had') implies that the hearer must remember an earlier mention of the subject. Heta's response (line 3) signals that she understands the references and agrees with Irja's plan.

The second *tuo*-form, the plural adessive form *noil* 'on those', refers to people, because it is produced as a self-repair in which the self-reference (*mul* 'on me') is substituted with it. The third form, the locative adverb *tuol* 'over there', refers to a place other than the speaker's location. The relevant portion of the utterance is to relate that the 'First-aid for dogs' is not in the speaker's possession, but that she can check for it later. The questions who has it, where it is, and whether it should be described as a book or a leaflet are not salient, particularly because the recipient can access the information herself by recollecting an earlier discussion. This follows the Gricean principles that a speaker must give only as much information as needed and no more and only say things that are relevant to the conversation.

How does this kind of vague use of *tuo* compare to other means of expressing vague, unclear, or indefinite referents in Finnish? The indefinite pronouns *joku* 'somebody' and

jokin 'something' imply that the referent is unknown, not only to the hearer but also to the speaker. These pronouns are also occasionally used to express indifference. In addition to these indefinite pronouns, Finnish has four pronouns that are described as specific indefinite. They are used when the speaker can identify a specific referent but presents it as unknown to the hearer, thereby implying that the hearer has no access to the referent. These pronouns are *eräs* and *muuan*, which imply 'certain', *yksi* 'one', and *tietty* 'known'. *Eräs* and *muuan* are rather formal in style. Occasionally, *yksi* 'one' is also used like an indefinite article in spoken language.

The demonstrative pronoun *tuo* differs from the above pronouns in terms of the kind of participant access it implies. As mentioned above, in (7), the speaker presents the non-salient referents, thereby implying that the recipient may access them independently. *Tuo*, which implies shared (un)accessibility, also aligns with Etelämäki's (2006) account of the semantic features of Finnish demonstrative pronouns.

References with *tuo* forms to non-explicated targets include a significant number of adverbial forms referring to 'somewhere over there', as in (7), which have an adessive case ending (*tuolla, tuol, tual* 'over there'[7]). Adverbial forms with the inessive case ending (*tuossa, tossa, tos*) are used vaguely to refer to a time, usually meaning something like 'some time ago', as in (8).

In (7), the locative adverb rooted to *tuo* expresses non-salience and vagueness. Similar uses of the adverbs expressing a time are also found in the data. In (8), sisters Tuula and Jaana are discussing getting old. Tuula mentions that their mother had jokingly reminded her that Tuula will be celebrating her fiftieth birthday next year. *Tossa* 'over there', in my interpretation, refers to a time when the reported conversation took place and translates as 'that time' or 'recently'. In the data, this kind of *tossa* reference may be accompanied by some other expression of a time—for example, *tossa viimeviikolla* ('last week').

(8) SG438

01 Tuula:	*Et*	*se*	*oli*	*äiki*	*äiti*	*naureskeli*	*mulle*	***tossa***	*että,*
	that	3SG	was	MS	mother	laughed	I.ABL	over.there	that

'So it was mother laughing at me recently that'

02 *.mth @kukas se täyttää ↑viiskymmentä ↑ens vuonna?@ >°Mä et°<*

03 *↑jo↓o? (.) todellaki mi↓nä nyt täytän viiskymmentä ens vuonna?*

'who is going to turn 50 next year? I was like "now already? 'Really I'm turning 50 next year?"'

Examples (7) and (8) have presented occurrences of *tuo* pronouns inflected in locative cases and used as adverbials. In contrast, the next example (9) demonstrates a nominative form in subject position. Repeating *tuo* expresses a difficulty in accessing the correct lexical term for the referent. In subject position, referents tend to be salient and usually obtain a lexical definition after word search, as in (9). In the example, the speakers are teenage girls who are doing their mathematics homework together.

[7] Numerous locative adverbial forms are identical with the inflected forms of the pronoun *tuo*, but some forms are separate: for example, *tuolla* means both 'on that' and 'over there', but the colloquial *tol* means only 'on that'.

(9) SG120
01 Milja: *Ku mie ↑en osaa näit yh:tää.*
'Cause I can't do these at all.'
02 Oona: ***Tota:**, [Milja, **toi** on, (.) **toi** **toi**, (.) **toi** o'*
well NAME that is that that that is
'Well, Milja, that is'
03 Milja: *[Mhy?*
PTC
04 Oona: *halkasija.h.*
'a diameter.'
05 Milja: *>Ai nii<?*
'Oh yes.'

The first partitive form of *tuota* in line 2 functions to draw the other participant's attention to the speaker's utterance. Milja's response is produced simultaneously with Oona saying her name. Then, Oona points out a part that Milja has misunderstood but by repeating the demonstrative pronoun in nominative; she expresses trouble in finding the word 'diameter'. Nominative subjects may be left unclear in a way that resembles the inflected *tuo*-forms in examples (7) and (8). For instance, this may happen in contexts where a non-explicated reference targets someone who is the original witness of an event that a speaker is reporting (see Priiki 2020: 195; see also example (10) further down in this article). Another context in which unclear referents occur is an abstract situation as the target of the reference. I examine occurrences of *tuo* in these contexts in the next section, as, in these contexts, the focus is on the ambiguity regarding whether or not the pronoun form is referential.

In this section, I have attempted to shed light on the question of why *tuo* 'that' is selected to pragmaticize to a hesitation particle. I have suggested that this can be explained partially by the type of participant access it implicates. As Etelämäki (2006, 2009) has shown, references with *tuo* indicate that the referent is equally accessible or non-accessible to the speaker and the recipient. Another feature is that *tuo* refers to targets that are unimportant and, thus, a vague reference is sufficient for the conversation.

5.2 Questionable referentiality

The occurrences of the pronoun *tuo* discussed in the previous section were all syntactic constituents or determiners, even though the target of their reference was unclear. In this section, I examine cases that are more advanced in their process of pragmaticizing to a particle. This implies that it is difficult to tell whether they are referential or whether their function is merely to delay completing the utterance or to fill a syntactic slot that needs to be occupied in a certain structure. In the previous section, I have shed light on the question of why *tuo* in particular is chosen to pragmaticize to a hesitation particle. In this section, I reflect on the possible contexts where the pragmaticization may have taken place. I suggest that the line between placeholder function in, for example, a tail construction and non-referential filler function is a fuzzy one and that this could be one factor in explaining how *tuota* acquired the hesitator function.

In Section 3, I noted that the partitive form *tuota* is more common as a determiner of a noun phrase than in other positions. Lappalainen (2010: 128) suggested that the determiner position would be the function where the pronoun has turned to a particle, and

the difference in frequency may give the same impression. In Arkisyn data, *tuo* determiners are used in various cases to express word search and hesitation, as presented above in examples (2b–c), (4), and (5). When *tuo* is inflected in other cases than partitive, it is more easily interpreted as referential. However, as already shown in example (3), non-referential nominative forms are also frequent in the data. In both nominative and partitive forms, there is a significant number of such cases where the interpretation of words *tuota* or *tuo* as a determiner of a noun may be questioned, as in the next example (10).

(10) SG438
01 *-- et ainoo mikä sitä nyt, (.) nytte=kää kiinnostaa ni on*
02 *se justii että: et se vaan tatu#oi itteensä#.*
'- - that the only (thing) that interests her now is to get herself tattooed.'
03 *(0.2)* ***Toi****, (.)* *äiti* *sano* *justii että, .hhhh*
that/well mother said just that BREATH
04 *et* *ei=ks* *se, (.) ↑satu* *jo* *hirveesti* *tommone - -*
that NEG=CLI 3SG hurt only horribly that.kind
'Well, mother just said that, ehm, won't it hurt a lot, that kind of (stuff) - -'

In (10), the speaker is wondering about her daughter's eagerness to get tattoos. She refers to her mother and quotes her words to support her attitude. The nominative form of *tuo* precedes the word 'mother', which is also in the singular nominative form. Without a slight pause between the words, it would more straightforwardly be interpreted as a determiner. A *tuo* determiner in this kind of context would be natural in spoken Finnish: it would signal that the referent, the mother of the speaker, has not been discussed before this and will not become a salient topic. It is the quote that is central for the flow of the conversation, not the person who is quoted. However, the pause makes it possible to interpret the *tuo* word as a non-referential particle, only expressing that the speaker is processing how to continue. The word 'mother' without any determiners would also be a natural option.

In this article, I have shown that not only the partitive form of the demonstrative pronoun *tuo* but also the nominative form are used in particle function. However, non-referential partitive forms are more frequent in conversational data. The partitive form is a frequent object case in Finnish. Referents presented as objects of a clause are often new information and, thus, more difficult to access. In the next example (11), there appear two partitive occurrences of *tuota*. Whether or not they are referential is questionable. The speakers, Missu and Vikke, are discussing what to buy for a present for a friend.

(11) SG112
01 Missu: ***Mitä*** *siltä* *puuttuu.*
what.PART 3SG.ELA lack
'What would she need?' (lit. 'What does she lack?')
02 Vikke: ***Tota****,* *(no)* *ku* *se* *Hanne=ki* *on* *vähän* *osta-nu*
that.PART well because the NAME=CLI is a.little buy-PCP
03 *semmos-ta su-* ***tota****,* *semmos-ta* *lehmä-sarja-a?*
such-PART that.PART such-PART cow-series-PART
'Well, Hanne, too, has bought her some (parts of) that, a kind of cow(-themed) set (of dishes).'

In (11), the partitive form *tuota* occurs at the beginning of an utterance. The turn is an answer to a question, 'what would she need?' in line 1, where the question word *mitä* 'what' is also in partitive form. *Tuota* projects an answer to this question. The partitive question word in the preceding question makes it possible to interpret *tuota* as referential, which would be the case if Vikke began her answer with a lexical description (e.g. *tota semmosta lehmäsarjaa*). As a particle, *tuota* can project an answer to any kind of a question, and, as Etelämäki and Jaakola (2009) describe, it precedes a phrase that is still being formulated. Vikke does not answer the question directly but begins to explain what another friend has bought. The partitive *tuota* holds a place for the answer and creates an anticipation of Vikke telling what the friend lacks. The partitive form is repeated in line 3, preceding the lexical description of the present Vikke is suggesting. The noun phrase is preceded by multiple self-repairs, where the partitive forms of a demonstrative adjective *semmosta* and *tuota* alternate. Before choosing the word *lehmäsarjaa* 'cow(-themed) set', the speaker attempts another noun phrase (*su-*). Self-repairs are another typical context where *tuota* particles are used (Lappalainen 2004: 128–131). In this context, *tuota* could project just the repair that is coming or it could be produced as a determiner, which is subsequently replaced by the demonstrative adjective.

In the data examined in this article, there are several occurrences of *tuota* at the beginning of an utterance where it may be possible to interpret it as a placeholder for the object of the thought or statement. The objects for such verbs are often abstract entities that are difficult to define with simple noun phrases. The referential relationship may appear unclear when the word projects a certain abstract line of thought. This kind of use may relate to Podlesskaya's (2010: 21) note that, occasionally, a non-default form of a hesitator demonstrative may create an impression of a certain elliptic verb in a structure that resembles the English placeholder *whatchamacallit* but without an explicit 'call' verb; the placeholder form is the object for the verb 'call'.

In (12), the speakers are discussing the timetable of a participant's planned graduation. In line 5, Iiro changes the topic a little, asking about the length of this participant's thesis. The word *tota* occurs at the beginning of his turn to speak.

(12) SG441
01 *Mari: .Hhy mutta siis jouluna kakstuhattaneljätoista.*
'But Christmas 2014.'
02 Jussi: *Teoriassa, teoriassa.*
'In theory, in theory.'
03 Elli: *↑Mm, (tai) keväällä kakstuhattaviistoist silloha*
04 *[se ois jo kuus vuotta (--),]*
'Or spring 2015 then it would be six years already - -'
05 Iiro: *[**Tota:** (mä rupesi miettiin) et] pitää=k se su diplomityän*
that.PART I started thinking that should=CLI the your thesis
olla joku kuuskyt sivuu jottain shaiba-a?
be some sixty page.PART some.PART rubbish-PART
'This/well I started wondering that should your thesis be like sixty pages of some rubbish?'

At first glance, the word order does not support the thought that the *tuota* in (12) would be a placeholder for the object. Objects are usually located after the finite verb; the neutral order would be *mä rupesi miettiin tota* 'I started thinking that [thing]'. Despite the

unusual word order, interpreting a pronoun-originating placeholder as an expletive object is not a unique idea. Already Rapola (1954) indicated that the particles *sitä* and *tuota* give the impression that they could be expletive objects—or placeholders for a description that is difficult to formulate.

As already mentioned, Finnish has two other particles that originate from partitive forms of demonstrative pronouns, *sitä* and *häntä*. Their pragmaticization process resembles that of *tuota,* since they lost the connection to the number and the case of the possible referent. The place where they occur in the sentence has some similarity to *tuota*, as well, even though *tuota* has more freedom. The pragmaticizing of the pronoun *tuo* may be a part of a more general tendency of Finnish pronouns—particularly their partitive forms—to turn to discourse particles. I discuss this possibility in the next section.

6 Partitive forms *tuota, sitä*, and *häntä* in particle function

The particle *sitä* is originally the partitive singular form of the demonstrative pronoun *se* 'it; 3SG'. As with *tuota*, *sitä* has lost its referentiality and its meaning is difficult to describe. In the Arkisyn corpus, *sitä* forms are systematically coded to demonstrative pronouns instead of particles, even though Finnish grammar (Hakulinen et al. 2004: § 827) mentions *sitä* as a particle. Although Hakulinen (1975: 26) mentions *sitä* as a spoken language feature, in Arkisyn, *sitä* forms are generally less frequent than *tuota* forms and fully pragmatized occurrences are rare. In spoken dialects, yet another pronoun, the personal pronoun *hän* 'he, she' is used as a particle (Laitinen 2005, Soikkeli 2013). In Arkisyn, no occurrences of *hän* in particle function are found, and the personal pronoun is little used in informal everyday conversations. Resembling *tuota* and *sitä*, this particle is often frozen in the partitive form, *häntä*; however, other frozen forms also occur—for example, the adessive form in the phrase *hällä väliä* 'who cares'.[8]

The few examples of the *sitä* particle found in Arkisyn represent its typical contexts (see Hakulinen 1975, 29; Vilkuna 1989: 143–144; Hakulinen et al. 2004: § 827). In (13), the *sitä* particle occurs in a zero-person construction. The conversation is between a hairdresser and customer. The customer is telling the hairdresser about an electric warmer on the roof of his house and the hairdresser is expressing doubt regarding its safety.

(13) SG108

01	*Joo*	***sitä***	*vaan*	*kuvittelee*	*et*	*jos*	*ränni-s*	*on*	*sähkö*
	yeah	3SG.PART	only	imagine	that	if	gutter-INE	is	electricity
02	*£et(h)ä*	*s(h)e*	*on*	*v(h)aa[rall(h)inen£*	*he he.*				
	that	3SG	is	dangerous					

'Yeah [you] just imagine that if there is electricity in a gutter, it's dangerous.'

In (13), the position of *sitä* is similar to *tuota* in (12): the possible placeholder is situated before the finite verb and both verbs 'think' (in 12) and 'imagine' (in 13) would require a partitive case for their objects. Thus, both pronouns could be placeholders for the object, an abstract thought. However, Holmberg and Nikanne (2002) argue that *sitä*

[8] The gloss for the phrase is *3SG-ADE matter* and its literal meaning would be elliptic 'it (does not) matter' or '(what does) it matter'.

has no connection to case and agreement. According to them, it is rather used as a pure expletive merely to fill the required position in clauses that lack a natural subject, such as a zero-person construction. All Arkisyn examples, though rare, show a certain connection between *sitä* and the partitive case. *Sitä* used in intransitive clauses, documented in the earlier studies (Holmberg & Nikanne 2002; Vilkuna 1989: 144–145), seem to be rare in contemporary everyday conversations, or at least it is not found in this corpus.

While earlier studies present *sitä* as obligatory in certain contexts, such contexts are not frequent in contemporary conversation data. *Sitä* could be used in (12) without significantly changing the meaning—but, unlike *tuota*, *sitä* would, in this context, be clearly referential. *Tuota* would not be quite natural in (13), since zero-person construction typically occurs with *sitä*. However, among the few examples of the *sitä* particle found in Arkisyn, a few occurrences could be changed to *tuota* or even to *häntä*. Example (14) is from the same conversation as (13) and it is possible to interpret this as a tail construction where *sitä* is coreferential to the noun *lomaa* 'vacation' in singular partitive.

(14) SG108 (H=hairdresser, C=customer)
01 C: *Hh meinaaksä nyt pitää l- lomaa sitte heinäkuussa (.) °vai miten,°*
'Do you intend to have a holiday in July, or when?'
02 H: *Hh kyllä mä: tota m (.) .hhh ↑mä en ihan oikein tiedä sitte että mhh*
03 *koska mä pitäsin mutta (.)*
'Yes I well- I don't really know when I would have but'
04 *kyllä=hä* ***sitä*** *täytyy vähä yrittää pitää loma-a - -*
PTC=CLI 3SG.PART must little try have holiday-PART
'of course (one) must try to have a little time off - -'

Vilkuna (1989: 145) suggests that the function of *sitä* in this kind of order, where there is a verb-initial constituent (*kyllähä* in 14), is to ensure that the constituent preceding *sitä* is interpreted as topicalized. The same result would be obtained if *sitä* in (14) were replaced with *tuota*. All three pronouns or particles project something on the turn that follows, thereby indicating different implications. *Sitä* implicates that the speaker would like to introduce the topic of having a holiday for further discussion, while *tuota* presents the topic in a non-salient manner, or it would be interpreted as a hesitator. *Häntä*, in contemporary everyday speech, would sound playful and archaic, since it is so rarely used. The functions of these particle-like forms have been studied in different data. According to Hakulinen (1975), *sitä* softens questions, marks the utterance as a discussion opening, and guides a hearer to seek a metaphorical interpretation. According to Laitinen (2005) and Soikkeli (2013), *hän* used in particle function is connected to the functions of the pronoun *hän* in dialects in general, where it is typically used in quotes when referring to the original speaker.[9]

As mentioned, example (14) may be interpreted as a tail construction. Moreover, Vilkuna (1989: 139–141) notes the possible connection of the *sitä* particle to the tail construction, but indicates that while the placeholder pronoun of a tail construction may be freely situated anywhere in the clause, the *sitä* particle is tied to the verb-initial theme or topic position. Despite being free in principle, analysing naturally occurring conversations indicate that the tail construction placeholders are usually also situated in the same position (Priiki 2020: 195–196). Considering *häntä*, Laitinen (2005: 102) brings up the tail

[9] See Laitinen 2002, 2005; Nau 2002; Priiki 2017.

construction as a context where questionably referential pronouns occur. I suggest that the pragmaticizing process of the *tuota* particle may have some connection to the same phenomenon.

Why do these three pronouns tend to be used as particles, particularly in their partitive forms? The referents of a partitive NP are usually less individuated, less central to communication, and less frequently mentioned again than referents of an NP in accusative or nominative cases (Helasvuo 1996: 28–30). In the object role, the referentiality of a pronoun may easily become unclear, particularly when the referent is an abstract entity—such as the target of speaking, thinking, or imagining—as in the examples above.

7 Conclusions

In this article, I examined the continuum of referential, vaguely or questionably referential, and particle-like occurrences of the Finnish demonstrative pronoun *tuo* 'that'. I focused on the forms that are open to interpretation on this continuum, aiming to shed light on the question of why the partitive form *tuota* in particular has been pragmaticized to a particle expressing hesitation and word search. The study complements the examination of the hesitation word *tuota* in conversation data that was initiated by Etelämäki and Jaakola (2009). Their article focused on occurrences of *tuota* in clearly particle function without taking a stance on the context where the pragmaticizing of the pronoun has occurred.

I showed that not only the partitive forms but also other case forms of the pronoun may be used without a clear referent. As already noted by Etelämäki and Jaakola (2009), the semantic features of the pronoun *tuo* make it the most suitable to express hesitation among all the Finnish demonstratives. *Tuo* forms as determiners and placeholders can project the type of the referent while the speaker is still processing the lexical definition. While the other demonstratives *se* and *tämä* implicate that the referent is already known or that it is central for the conversation, *tuo* projects a non-salient referent that is only just becoming the target of attention. In spatial contexts, *tuo* is distal; in numerous other languages as well, distal demonstratives are selected for the placeholder and hesitator functions. What is peculiar in the consideration of the Finnish hesitator demonstrative is that the form pragmaticized to the particle function is the inflected singular partitive, *tuota*, while usually the singular nominative form is the most likely to lose the connection to case and number.

Analysing the borderline cases between the referential and non-referential functions of *tuo* reveals that numerous occurrences still retain referentiality, projecting subject, object or location, even though the exact referent is not explicated. The referentiality becomes questionable most easily when the referent is an abstract entity that cannot be lexicalized with a simple noun phrase, such as the object for speaking or thinking. These kinds of verbs usually have their objects in partitive case. In these cases, the partitive *tuota* can have an ambiguous interpretation of projecting either the object or the whole utterance. Numerous borderline cases resemble tail constructions where a referent is referred to twice—first with a demonstrative pronoun placeholder and then with a lexical noun phrase.

Further, I compared the particle *tuota* to two other Finnish particles that have pragmaticized from a partitive form of a demonstrative: *sitä* from *se* 'it; 3SG' and *häntä* from *hän* 'he, she'. I suggest that the process of *tuo* turning into a particle form is part of a more general tendency in Finnish for partitive forms to lose their referentiality when they are

used as placeholders for objects. In this function, they occur at the beginning of an utterance, thereby not only projecting implications regarding the possible object referent but also about the entire turn. The different pronouns that participate in this kind of pragmaticizing process still retain meaning features typical to the original pronouns, and the different meanings of the pronouns reflect the different functions of the particles.

Studying naturally occurring conversations can direct research to phenomena that are frequent in everyday speech but have not been thoroughly studied. The *tuota* particle has been little studied compared to the Finnish expletive *sitä*, even though the former is far more frequent in everyday speech. This study has revealed that the neglected and disapproved hesitation particle *tuota* is, in fact, a rather complex phenomenon. Understanding its behavior would require further study—for example, focusing on prosodic patterns and the particle chains it tends to form.

Data sources

Arkisyn Database of Finnish Conversational Discourse. Compiled at the University of Turku, with material from the Conversation Analysis Archive at the University of Helsinki and the Syntax Archives at the University of Turku. Department of Finnish and Finno-Ugric Languages, University of Turku. Available via Kielipankki, the Language Bank of Finland at http://urn.fi/urn:nbn:fi:lb-2017022702

References

Etelämäki, Marja. 2006. *Toiminta ja tarkoite. Tutkimus suomen pronominista tämä* [Activity and referent: A study of the Finnish pronoun *tämä*]. Helsinki: Suomalaisen Kirjallisuuden Seura [The Finnish Literature Society].

Etelämäki, Marja. 2009. The Finnish demonstrative pronouns in light of interaction. *Journal of Pragmatics* 41. 25–46. https://doi.org/10.1016/j.pragma.2008.09.005

Etelämäki, Marja & Jaakola, Minna. 2009. *Tota* ja puhetilanteen todellisuus ["Tota" and the reality of speech situation]. *Virittäjä* 113. 188–212.

Frank-Job, Barbara. 2006. A dynamic-interactional approach to discourse markers. In Kerstin Fischer (ed.), *Approaches to discourse particles*, 395–413. Amsterdam: Elsevier.

Ford, Cecilia. 1993. *Grammar in interaction*. Cambridge: Cambridge University Press.

Hakulinen, Auli. 1975. Suomen *sitä*: Pragmatiikan heijastuma syntaksissa [The syntax and pragmatics of Finnish *sitä*]. *Sananjalka* 17. 25–41. https://doi.org/10.30673/sja.86392

Hakulinen, Auli & Seppänen, Eeva-Leena. 1992. Finnish *kato*: From verb to particle. *Journal of Pragmatics* 18, 527–549. https://doi.org/10.1016/0378-2166(92)90118-U

Hakulinen Auli, Vilkuna, Maria, Korhonen, Riitta, Koivisto, Vesa, Heinonen Tarja Riitta & Alho, Irja. 2004. *Iso suomen kielioppi* [The big grammar of Finnish]. Helsinki: Suomalaisen Kirjallisuuden Seura [The Finnish Literature Society]. http://scripta.kotus.fi/visk

Hayashi, Makoto & Yoon, Kyung-Eun. 2006. A cross-linguistic exploration of demonstratives in interaction. With particular reference to the context of word-formulation trouble. *Studies in Language* 30(3), 485–540. https://doi.org/10.1075/sl.30.3.02hay

Helasvuo, Marja-Liisa. 1996. A discourse perspective on the grammaticization of the partitive case in Finnish. *SKY 1996 Yearbook of the Linguistic Association of Finland*, 7–34.

Heine, Bernd. 2013. On discourse markers: Grammaticalization, pragmaticalization, or something else? *Linguistics* 51(6). 1205–1247. https://doi.org/10.1515/ling-2013-0048

Holmberg, Anders & Nikanne, Urpo. 2002. Expletives, subjects and topics in Finnish. In Peter Svenonius (ed.), *Subjects, expletives, and the EPP,* 71–105. Oxford: Oxford University Press.

Huumo, Tuomas. 2010. Nominal aspect, quantity, and time: The case of the Finnish object. *Journal of Linguistics* 46. 83–125. https://doi.org/10.1017/S0022226709990223

Keevallik, Leelo. 2010. The interactional profile of a placeholder. The Estonian demonstrative 'see'. In Nino Amiridze, Boyd Davis & Margaret Maclagan (eds.), *Fillers, pauses, and placeholders.* (Typological studies in language 93.), 139–172. Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.93.07kee

Laitinen, Lea. 2002. From logophoric pronoun to discourse particle: A case study of Finnish and Saami. In Ilse Wischer & Gabriele Diewald (eds.), *New reflections on grammaticalization*, 327–344. Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.49.21lai

Laitinen, Lea. 2005. *Hän*, the third speech act pronoun in Finnish. In Ritva Laury (ed.), *Minimal reference. The use of pronouns in Finnish and Estonian discourse*, 75–106. Helsinki: Suomalaisen Kirjallisuuden Seura [The Finnish Literature Society].

Lappalainen, Hanna. 2004. *Variaatio ja sen funktiot. Erään sosiaalisen verkoston jäsenten kielellisen variaation ja vuorovaikutuksen tarkastelua* [Variation and it's functions. The analysis of linguistic variation and interaction among members of a social network]. Helsinki: Suomalaisen Kirjallisuuden Seura [The Finnish Literature Society].

Larjavaara, Matti. 1990. *Suomen deiksis* [Finnish deixis]. Helsinki: Suomalaisen Kirjallisuuden Seura [The Finnish Literature Society].

Larjavaara, Matti. 2019. *Partitiivin valinta* [Choice of partitive]. Helsinki: Suomalaisen Kirjallisuuden Seura [The Finnish Literature Society].

Latvala, Salu. 1894. Lauseopillisia havaintoja Luoteis-Satakunnan kansankielestä [Observations about the syntax of the Northwest Satakunta vernacular]. *Suomi* III, 12.

Laury, Ritva. 1997. *Demonstratives in interaction: the emergence of a definite article in Finnish.* Amsterdam: John Benjamins.

Lönnbohm, O. A. F. 1879. Jääsken, Kirvun ja osittain Rautjärven ja Ruokolahden pitäjien kielimurteesta [About the dialects of Jääski, Kirvu and partly Rautjärvi and Ruokolahti parishes]. *Suomi* II, 13, 1–163.

Penttilä, Aarni. 1963. *Suomen kielioppi* [The Finnish grammar]. 2nd edition. Porvoo: WSOY.

Pihlaja, Aila. 1971. *Täytelisäys, puheen "parasiitti"* [Filler word, a parasite of speech]. Master's thesis, University of Jyväskylä.

Priiki, Katri. 2015. *Se oli iha hullu se hammaslääkäri.* Kaakkois-Satakunnan henkilöviitteiset *se, hän, tää* ja *toi* eteenpäin lohkeavan konstruktion osina ja ensimainintoina [Third-person pronouns *se, hän, tää* and *toi* in right-dislocation constructions and as a first mention of a person]. *Puhe ja kieli* 35(2). 47–71.

Priiki, Katri. 2020. The Finnish tail construction as first mention. *Nordic Journal of Linguistics* 43(2). 181–203. https://doi.org/10.1017/S0332586520000104

Podlesskaya, Vera. 2010. Parameters for typological variation of placeholders. In Nino Amiridze, Boyd Davis & Margaret Maclagan (eds.), *Fillers, pauses, and placeholders.* (Typological Studies in Language 93), 11–32. Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.93

Ravila, Paavo. 1945. Lauseeseen liittyneet irralliset ainekset [The loose parts of a sentence]. *Virittäjä* 49. 1–16.

Rapola, Martti. 1954. Erään tyylipiirteen taustaa [The background of a stylistic feature]. *Virittäjä* 58. 1–17.

Reile, Maria, Piia, Taremaa, Nahkola, Tiina & Pajusalu, Renate. 2019. Reference in the borderline of space and discourse: A free production experiment in Estonian, Finnish and Russian. *Linguistica Uralica* 55(3). 185–208. https://dx.doi.org/10.3176/lu.2019.3.02

Setälä, Emil Nestor. 1883. Lauseopillinen tutkimus Koillis-Satakunnan kansankielestä [A syntactic study of the Northeast Satakunta vernacular]. *Suomi* II: 16.

Soikkeli, Katja. 2013. *Elotonviitteinen ja partikkelistunut* hän-*sana ja sen käyttö 1960-luvun murrehaastatteluissa. mitä hänessä lie* [*Hän* with inanimate referents and as a particle in 1960s dialect interviews]. Master's thesis, University of Helsinki.

Vilkuna, Maria. 1989. *Free word order in Finnish. Its syntax and discourse functions.* Helsinki: Suomalaisen Kirjallisuuden Seura [The Finnish Literature Society].

Vuorinen, Riitta. 1981. Puhekielen täytelisäkkeistä [About filler words in spoken language]. In Matti Kalevi Suojanen (ed.), *Kirjoituksia puhekielestä. Turun puhekielen projektin julkaisuja 1.* [Writings about spoken language], 79–96. Turku: The Department of Finnish and General Linguistics, University of Turku.

Wide, Camilla. 2011. Pronomen *den här* och *den där* som planeringspartiklar i finlandssvenska dialekter [Pronouns *den här* och *den där* as planning particles in Finland Swedish dialects]. In Gustav Bockgärd & Jenny Nilsson (eds.), *Interaktionell dialektologi* [Interactional dialectology], 251–306. Uppsala: The Institute for Language and Folklore.

Östman, Jan-Ola. 1995. Recasting the deictic foundation using physics and Finnish. In Masayoshi Shibatani & Sandra Thompson (eds.), *Essays in semantics and pragmatics: in honor of Charles J. Fillmore*, 247–298. Amsterdam: John Benjamins. https://doi.org/10.1075/pbns.32.13ost

Katri Priiki
School of Languages and Translation Studies, University of Turku
katri.priiki@utu.fi

Object Marking with Discrete Objects in Finnish and Lithuanian

Asta Laugalienė

Abstract: The case of the direct object of transitive verbs in Finnish alternates between the accusative and the partitive. In Lithuanian, there is an alternation between the accusative and the partitive genitive. It was shown in previous research that some functions of the Finnish partitive and the Lithuanian partitive genitive in object marking are identical (i.e. partial affectedness of mass nouns) but there are some features that haven't received enough attention in the literature, e.g., the Lithuanian partitive genitive with discrete nouns. This paper offers an overview of possible conditions for the use of partitive genitive in resultative constructions in modern and older Lithuanian in comparison with their counterparts in Finnish and Slavic.

Keywords: *irresultativity, partitivity, object marking, partitive, Finnish, Lithuanian*

1 Introduction[1]

The topic of this paper is irresultative meaning in object marking in Finnish and Lithuanian and their neighbours. The resultative situation is interpreted in this article as a transition in which the event brings about a change, after which there is no return to the original state but entry into another one. Its opposite, the irresultative situation, implies that no such transition takes place and after completion of the event the situation returns to the original state or to a state that is conceptualized as similar to the original state. Both Finnish and Lithuanian make the resultative *versus* irresultative distinction in object marking.

The resultative and irresultative readings of some Finnish achievement verbs, such as *ampua* 'shoot' have been discussed at length by many authors (for example, Heinämäki 1984: 153, Kiparsky 1998: 2–3). It is stated that the resultative (1a) *versus* irresultative opposition (1b) indicates the achievement or absence of a result:

(1) a. *Ammuin karhun.*
shoot.PST.1SG[2] bear.ACC.SG
'I shot the (a) bear.'
b. *Ammuin karhua.*
shoot.PST.1SG bear.PAR.SG
'I shot at the (a) bear (without killing it).' (Finnish, Kiparsky 1998: 2–3)

In Lithuanian, the irresultative use of the partitive genitive seems to be very rare. Many scholars (for example, Larsson 1983: 135, Koptjevskaja-Tamm & Wälchli 2001: 654, Seržant 2014: 286, Seržant 2015: 389) mention the fact that in Eastern Lithuanian dialects the partitive genitive may be used instead of the accusative in order to encode the temporariness of the result of a transfer (2a–b). The accusative object in (2c) has no implications of temporariness and is used in standard Lithuanian:

[1] I wish to thank two anonymous reviewers as well as Axel Holvoet for their critical reading and valuable comments, which helped to improve the article.

[2] Abbreviations: 1 – first person, 2 – second person, 3 – third person, ACC – accusative, ABL – ablative, AOR – aorist, ART – article, COM –comitative, COMP – comparative, CVB – converb, DAT – dative, ELA – elative, F – feminine, GEN – genitive, ILL – illative, IMP – imperative, INE – inessive, INF – infinitive, INS – instrumental, LOC – locative, M – masculine, MED – middle voice, NEG – negation, NOM – nominative, PAR – partitive, PL – plural, POSS – possessive, PP – passive participle, PPA – past participle active, PRS – present, PST – past, PVB – preverb, Q – question particle, RFL – reflexive, SG – singular, TR – translative.

(2) a. *Duok man peilio!* (neilgam, tuoj sugrąžinsiu)
give.IMP me knife.GEN.SG
'Give (me) a/the knife! (for a while, I will shortly give it back to you)'
(Lithuanian, Jablonskis 1957: 578)

b. *Paskolink peilio!*
PVB.lend.IMP knife.GEN.SG
'Lend (me) a/the knife!' (Lithuanian, Ambrazas et al. 1976: 25)

c. *Duok peilį!*
give.IMP knife.ACC.SG
'Give (me) a/the knife!' (Lithuanian, personal knowledge)

Interestingly, both examples with the partitive genitive (2a–b) come from the same two sources (Jablonskis 1957: 578 (2a) and Ambrazas et al. 1976: 25 (2b)) and are cited again and again by scholars. Moreover, in Ambrazas et al. (1976) there is a reference to the example given by Jablonskis (1957: 578), which is basically the same as the one cited by Ambrazas et al. (1976). In the case of the author of this article, neither her intuition as a native speaker of modern Lithuanian nor her own dialect (Northern Lithuanian) suggest that the use of the partitive genitive in such sentences would be possible. This observation was one of the starting points for this research, which aims to answer the question in which situations partitive or partitive genitive is interpreted as encoding an irresultative event in Lithuanian compared to other neighbouring languages. The main focus will be on Lithuanian and Finnish, the latter – as will be shown – having much wider criteria to encode irresultative events than Lithuanian.

The idea behind the present article was to bring a new perspective to the widely investigated research domain of Fennic and Baltic object marking by concentrating on object marking with discrete objects as a separate topic, but also by using Lithuanian diachronic and dialectal data in order to show that irresultative partitive marking must once have been more widespread in both Fennic and Baltic, though standard Lithuanian has almost completely lost it.

The goal of this paper is to describe the semantic factors that give rise to the variation in the case-marking of discrete objects in Finnish, Lithuanian and the neighbouring languages (Polish, Russian, Estonian etc.), with a comparison between Lithuanian and Finnish in the foreground. Another research question, which the present article aims to answer, is to confirm the hypothesis, that Lithuanian and Finnish might have different strategies for encoding irresultativity in discrete objects (aspectual prefixes vs. partitive marking). If so, no consistent marking of irresultativity via case-marking would be expected in Lithuanian.

Examples for this research are taken from various sources, each of them marked separately next to the example. Old Lithuanian was checked against the old Lithuanian Corpus.[3] In the absence of electronic corpora of dialectal Lithuanian of all regions, the dialectal data was checked against the sources, which were available at hand, also some informants were consulted.

The following Section 2 provides the background of the study on object case marking in Baltic, Fennic and Slavic. Section 3 discusses the semantic classification of irresultative constructions and gives further observations. Sections 4 and 5 are devoted to discussion and concluding remarks.

[3] The Old Lithuanian corpus contains texts from the 16th to the 20th centuries, each century is represented by about 1 mln words. A list of verbs, which could be expected to have partitive genitive with discrete objects was drawn up based on the occurrences of partitive objects in neighbouring languages. Both prefixed and non-prefixed verbs were checked against the corpus in question. For more explanations about the data see Section 3.

2 Object case marking in Baltic, Fennic and Slavic

In Finnish, the case of the direct object of transitive verbs alternates between the 'total object' (marked with the accusative) and the 'partial object' (morphologically marked with the partitive). In this article, the term 'accusative' will be used as a blanket term for the non-partitive case forms.[4] The three interrelated and often overlapping functions of the partitive case in Finnish are: quantitative unboundedness of the object referent, which often correlates with an indefinite reading (3); aspectual unboundedness or lack of culmination in the designated event (4); and negation of the propositional content (5). They have been widely discussed in literature (e.g. Vainikka & Maling 1996: 193, Koptjevskaja-Tamm & Wälchli 2001: 650–652, Laugalienė 2020):

(3) *Löysin* *marjoja.*
find.PST.1SG berry.PAR.PL
'I found [some] berries.'

(4) *Kuuntelin* *radiota.*
listen. PST.1SG radio.PAR.SG
'I was listening to the radio.'

(5) *En* *rakentanut* *taloa.*
NEG build.PST.PA house.PAR.SG
i. 'I did not build a/the house.'
ii. 'I was not building a/the house.' (Finnish, personal knowledge)

In a nutshell, the multifunctional nature of the alternation between Finnish total and partial objects could be described as follows (Larjavaara 2019: 199): the object of the sentence is **total** whenever and only when a positive sentence expresses a complete change of the event that has reached (or is reaching) its endpoint (6).[5] In all other cases (including transitive sentences denoting some extent of change, e.g. *lämmitin saunaa*-PAR 'I heated the sauna (a bit)' or no change at all, e.g. *katsoin televisiota*-PAR 'I was watching TV'), the **partial** object is used as in (7):

(6) *Rakensimme* *talon.*
build.PST.1PL house.ACC.SG
'We built a house.'

(7) *Rakennamme* *taloa.*
build.PRS.1PL house.PAR.SG
'We are building a/the house.' (Finnish, personal knowledge)

In Baltic and Slavic, it is the genitive that most closely resembles the Finnish partitive (Koptjevskaja-Tamm & Wälchli 2001: 652). However, the use of the Lithuanian partitive genitive differs from that of the Finnish partitive in many respects. The most common use of the Lithuanian

[4] This case has very little dedicated morphology and is thus largely a non-autonomous case which borrows forms from other cases (on the notion of non-autonomous case, see Blake 2004: 22–24). For singular NPs, the object marker *-n* is homophonous with genitive case; plural direct objects are marked with the nominative plural. A dedicated form (the *-t* accusative) is used for personal pronouns, for example *he* 'they.PL.NOM': *heidät* 'they.PL.ACC'. The accusative case is thus defined mainly on the basis of syntactic context.

[5] Negation logically falls under this condition as the propositional content of the sentence is negated, which means that there was no culmination of the event (for more details on negation see e.g. Miestamo 2014: 67–70 or ISK 2004: § 932). Same applies to the sentences, where the actuality of the propositional content is doubtful, e.g. *Tuskin Jukka on lukenut kirjaa-PAR* 'It is unlikely that Jukka has read a/the book'.

genitive is with indefinite non-incremental quantification (where the genitive is used to refer to an indefinite number or quantity):

(8) *Jis rado knygų.*
3SG find.PST.3SG book.GEN.PL
'He found some books.' (Lithuanian, personal knowledge)

In Lithuanian accusative is used in the case of incremental quantification, when the object participates in the event in an incremental, gradual way, and its components are affected sequentially (9), but also for definite mass nouns (10), and in generic sentences (11) (Laugalienė 2020):

(9) *Aš geriu kavą.*
I drink.PRS.1SG coffee.ACC.SG
'I am drinking coffee.'

(10) *Išgėriau kavą.*
PVB.drink.PST.1SG coffee.ACC.SG
'I drank up the coffee.'

(11) *Geriu tik kavą.*
drink. PRS.1SG only coffee. ACC.SG
'I drink only coffee.' (Lithuanian, personal knowledge)

In Lithuanian, in line with Finnish (5), the direct objects of transitive verbs (even those normally marked with the accusative) will take the genitive case in negated clauses; this is the so-called genitive of negation, which historically evolved from the partitive genitive (Ambrazas 1997: 500–506, 667–668, see also Kuryłowicz 1971 for the Slavic genitive of negation):

(12) *Brolis nenusipirko naujo namo.*
brother NEG.PVB.RFL.buy.PST.3SG new.GEN.SG house.GEN.SG
'[My] brother did not buy a new house.' (Lithuanian, personal knowledge)

Aspectual distinctions in Lithuanian are often expressed by the choice of verbal prefixes (Koptjevskaja-Tamm & Wälchli 2001: 652). The difference in aspect is marked in imperfective/perfective verbs as in examples (13) and (14), but the case marking is not in itself a device used to differentiate aspect. The partitive genitive is mostly possible only with perfective verbs:

(13) *Pa-rašiau laišką.*
PVB-write.PST.1SG letter.ACC.SG
'I wrote a/the letter.'

(14) *Rašiau laišką.*
write.PST.1SG letter.ACC.SG
'I was writing a/the letter.' (Lithuanian, personal knowledge)

In Polish, as in Lithuanian, direct objects are encoded by genitive in negated clauses. Genitive objects refer to quantitatively unbounded entities almost exclusively in the context of perfective verbs, therefore aspect in Polish is relevant for the occurrence of partial objects (Koptjevskaja-Tamm & Wälchli 2001: 653). In Finnish – as shown in examples (3) and (4) – both indefinite

quantity and imperfectivity can, independently of each other, trigger partitive marking on objects. Thus both the genitive in Polish and partitive in Finnish are sensitive to aspect, but Finnish and Polish systems take completely opposite directions with respect to object marking for imperfective and perfective clauses: Finnish partitive is favoured by imperfective contexts and Polish genitive is favoured by perfective contexts (Koptjevskaja-Tamm & Wälchli 2001: 653–654). In Finnish, an imperfective context automatically leads to the partitive marking of the object whereas in Polish aspectual characteristics provide an additional restriction on the occurrence of the genitive object (ibid.).

Even though the alternation between total and partial objects is well-known from some of the older Indo-European languages (Brugmann & Delbrück 1897–1990: 575ff, cited by Koptjevskaja-Tamm & Wälchli 2001: 663), aspectual considerations are not mentioned as influencing the choice between the two cases at this stage. They appear as a factor in both Fennic and Balto-Slavic, but the developments were separate and led to different results.

At a first glance, there are considerable differences in the types of entities that could be treated as partial objects in Finnish, Lithuanian and Polish (Slavic). For Finnish mass nouns, the basic function of the partitive object is either non-culminating aspect or bounded non-specific quantity (or both):

(15) *Join* *kahvia.*
drink.PST.1SG coffee.PAR.SG
(i) 'I was drinking (the) coffee.'
(ii) 'I drank some (of the) coffee.' (Finnish, personal knowledge)

In (16a) *kirje* 'letter' designates a quantitively bounded discrete entity, and the action has not culminated in a result (either only a part of the letter was written or the process of the writing of the letter is still ongoing). The discrete object remains undivided but the activity covers only its parts. In (16a), the partitive appears only when part of the letter was affected by the event of the writing. The difference with respect to total affectedness follows from the fact that the whole entity was not targeted and the rest remains unaffected (Luraghi & Kittilä 2014: 41):

(16) a. *Kirjoitin* *kirjettä.*
write.PST.1SG letter.PAR.SG
(i) 'I wrote some of the letter.'
(ii) 'I was writing a/the letter.'
b. *Kirjoitin* *kirjeen.*
write.PST.1SG letter.ACC.SG
'I wrote a/the letter.' (Finnish, personal knowledge)

With respect to the marking of partially affected discrete objects, Lithuanian is different from Finnish, which marks partial affectedness via case. Partial affectedness of the discrete object in Lithuanian is encoded not in the object marking (both partially and fully affected objects are marked with the accusative), but in different prefixes of the verb, see (17a) vs. (17b):

(17) a. *Pa-skaičiau* *knygą.*
PVB-read.PST.1SG book.ACC.SG
'I read some of the book.'
b. *Per-skaičiau* *knygą.*
PVB-read.PST.1SG book.ACC.SG
'I read a/the book.' (Lithuanian, personal knowledge)

To conclude, Finnish positive clauses allow partitive object marking for discrete entities. Neither Lithuanian nor Polish (or Russian) normally allow discrete entities in affirmative positive clauses to be marked with genitive. Some exceptions to this rule will be discussed in Section 3.

3 Semantic classification

The collection of the data for this research was firstly obtained from various sources from Slavic and Fennic in order to investigate the possibilities for discrete objects to be marked with partitive genitive or partitive. Based on this research, lists of verbs expected to license partitive object marking with discrete objects were drawn up. These lists were checked against the Old Lithuanian corpus and also against available Lithuanian dialectical data in order to check whether and how partitive genitive marking with discrete objects is (or was) possible. Even though the examples from Lithuanian sources are not very numerous, the results show clear traces of such partitive genitive uses with discrete objects both in old Lithuanian and dialects.

Further below I suggest a classification of the semantic factors that give rise to the variation of the case marking of discrete objects in the languages examined. The classification is based on verbs which normally assign accusative to discrete objects and with which the use of partitive or partitive genitive is rather exceptional. The focus stays on the Lithuanian data, but other neighbouring languages are also taken into account. Based on the areal data (Baltic, Slavic and Fennic languages), four semantic groups could be established: temporally restricted usage, surface-contact verbs, scalar verbs and conative verbs. The Lithuanian data shows that temporally restricted usage is attested both in old Lithuanian and dialectical examples. Surface-impact and scalar verbs are not very well attested in old Lithuanian (there are no traces in the dialects), whereas the conative type is not attested at all (see table 1).

Semantic group	Verb	Translation	Number of occurrences	Source
Temporal restricted usage[6]	*skolinti*	lend	1	Ambrazas et al. 1976
	duoti	give	1	Jablonskis 1957
	užimti	take	1	URB 2013
	regėti	see	2	LT_16
	pamatyti	see	1	LT_20
Surface-contact verbs[7]	*prigriebti*	catch	1	LT_18
Scalar verbs[8]	*sudaužyti*	break	1	LT_19
Conative verbs[9]	–	–	–	–

Table 1: *Occurrences of verbs with partitive genitive for discrete objects in Lithuanian corpora*

3.1 Temporally restricted usage

In Lithuanian discrete entities can be marked with partitive genitive with certain verbs when there is a need to emphasize that the corresponding referents are given in someone's possession "for a little while" (Koptjevskaja-Tamm & Wälchli 2001: 654). A series of verbs like 'give' can take genitive, if the object is to be given for a limited time, i.e. some verbs allow two readings differing in the temporal stability of the resultant state.

Such a type of partitive genitive object seems to survive in Eastern Lithuanian dialects, where it has the implication of temporariness of the results of the transfer. Next to the example (2b), already cited in the introduction, the informants confirm more examples:

(18) *Duok kirvio!*
give.IMP axe.GEN.SG
'Give (me) an/the axe!' (Eastern Lithuanian, p.c. V. Kardelis)

Example (19) is taken from a book written in a local dialect of the Ukmergė region. Two informants confirmed that such use of the genitive object is normal in situations where the discrete object is given in someone's possession for a certain limit of time. In (19) *užimti kieno nors posto* means 'stand in for somebody' and the situation describes a temporary situation in which one person stands in for another:

6 Keywords *turėti* 'have', *daryti* 'open', *gauti* 'get', *padėti* 'put', *paguldyti* 'lay down', *nunešti* 'take', *pastatyti* 'put', *palikti* 'leave', *pririšti* 'tighten up', *išleisti* 'let out' yielded 0 results in Old Lithuanian corpus.

7 Keywords *plauti* 'wash', *tepti* 'spread', *valyti* 'clean', *remti* 'back up', *traukti* 'pull' gave 0 results in Old Lithuanian corpus.

8 Keywords *gadinti* 'spoil', *kelti* 'lift', *stabdyti* 'stop' gave 0 results in Old Lithuanian corpus.

9 Keywords *įrodinėti* 'argue, try to prove' *įkalbinėti* 'try to persuade' gave 0 results in Old Lithuanian corpus.

(19) *Po vieną bijodavom užimti jo posto,*
by one.ACC.SG be-afraid.PST.1PL occupy.INF he.GEN.SG post.GEN.SG
<kad nepraganytume karvių- dviese vis drąsiau.>
<so the cows wouldn't go astray - we were braver when there were two of us>
'Each on our own we were afraid to take his post, <so the cows wouldn't go astray - we were braver when there were two of us.>' (Lithuanian, URB 2013: 12)

In addition, it is attested both for older Russian (Koptjevskaja-Tamm & Wälchli 2001: 655) and Polish (Kempf 1970: 90), that there are certain verbs of perception or cognition that regularly combine with genitive. One of such verbs would be *regėti* 'see', which also takes the genitive object in Old Lithuanian as in (20) and (21):[10]

(20) *<Herodas Iesu ischwidens didei prassidʒuga nesa> iau senei isigeide*
<Herod was very happy to see Jesus, because> already long time want.PST.3SG
io regeti.
he.GEN.SG see.INF
'<Herod was very happy to see Jesus, because> he had been wanting to see him for a long time.' (Lithuanian, corpus LT_16)

(21) *Ir iéßkoio regét' Iésaus*
and look-for.PST.3SG see.INF Jesus.GEN.SG
<kas bût ir ne gałéio vʒ miniós nes' búwo mâʒo augléus>.
<and he was not able to see him amid the crowd as he was short of stature>
'And he sought to see Jesus <and he was not able to see him amid the crowd as he was short of stature>.'

These two examples are taken from the 16th century Biblical texts. This partitive genitive would be absolutely unusual for modern Lithuanian, which would have the accusative as in (23). In both (20) and (21) the genitive object could have been used to refer to a restricted time span, so that the meaning could have been 'cast a glance'. It seems that this temporally restricted usage of the genitive object has survived until the 20th century, as in (22) (in contrast with (23), which has the more frequent accusative):

(22) *<Prieangyje laukia moteris su mažyčiais verksniais kūdikėliais, mergaitės, atėjusios savo mylimųjų*
<In the entrance hall, there is a woman waiting with small crying babies, girls who came
aplankyti, ir> vyrai, norį pamatyti draugų,
to see their beloved ones and> man.NOM.PL want.PAP see.INF friend.GEN.PL
brolių ir savo žmonų.
brother.GEN.PL and own wife.GEN.PL
' <At the entrance hall, there is a woman waiting with small crying babies, girls who came to see their beloved ones and> men wishing to see [their] friends, brothers and wives.' (Lithuanian, corpus LT_20)

(23) *Portugalas Lietuvoje labiausiai norėjo pamatyti draugus.*
Portuguese.NOM Lithuania.LOC.SG most want.PST.3SG see.INF friend.ACC.PL
'A Portuguese man wanted most of all to see friends in Lithuania.' (Lithuanian[11])

[10] Animate objects regularly assume genitive marking in Russian and other Slavic languages. It is impossible to say whether animacy plays a role in Lithuanian as the old Lithuanian corpus did not give any results with inanimate objects.

[11] https://www.delfi.lt/verslas/verslas/cepelinai-ir-saltibarsciai-uzsieniecius-i-lietuva-vilioja-labiau-nei-merginos-ar-krepsinis.d?id=50816602

In Finnish, a series of verbs like *lainata* 'borrow, lend' can take partitive, if the object is given for a limited amount of time. Depending on the speaker's implications, both partitive and accusative are possible. In (24a) the girl is expecting to get her watch back in a while, whereas (24b) does not have such an implication:

(24) a. *Tyttö* *lainasi* *kelloa.*
girl.NOM lend.PST.3SG watch.PAR.SG
'The girl lent [her] watch [to somebody for a while].'
b. *Tyttö* *lainasi* *kellon.*
girl.NOM lend.PST.3SG watch.ACC.SG
'The girl lent [her] watch [to somebody].' (Finnish, personal knowledge)

Irresultative marking in Finnish applies to situations where the original state is almost the same as the target state. In (25a) the man raises his hat for a moment and puts it back: the target state does not significantly differ from the original state and also expresses the temporally restricted effect of the event. The sentence (25b), on the contrary, would indicate a transition from one state to another:

(25) a. *Mies* *nosti* *hattua.*
man.NOM raise.PST.3SG hat.PAR.SG
'The man raised (his) hat.' (Finnish, Leino 1991: 171–172)
b. *Mies* *nosti* *hatun* *päästään.*
man.NOM raise.PST.3SG hat.ACC.SG head.ELA.POSS.3SG
'The man took off his hat.' (Finnish, personal knowledge)

In Old Polish, the concept of the temporal partiality was very strong and the partial genitive instead of the accusative appear very consistently here. A series of verbs was oriented towards action limited in time, especially such as *dobyć* 'draw forth', *poprosić* 'ask', *(za)wołać* 'call', *udzielać* 'grant', *pożyczyć* 'borrow', e.g. *pożyczyć książki*.GEN 'give someone a book for a while, let him use it' (Kempf 1970: 192). However, genitive has remained productive in modern Polish with the verbs *dać* 'give' and *pożyczyć* 'lend'.

(26) *Daj* *mi* *ołówka.*
give.IMP me pencil.GEN.SG
'Hand me a pencil (for a while).' (Polish, Holvoet 1991: 110)

Verbs like 'give' can also take the genitive object in clauses with the meaning of temporal restricted use in Russian and Ukrainian (for Russian Buslaev 1959: 461, Kempf 1970: 190, for Ukrainian Shevelov 1963: 167, cited by Holvoet 1991: 110). In Northern Russian typical verbs are 'take', 'get', 'send', 'ask for' etc.

(27) *Voz'mu* *topora* *u* *vas.*
take.FUT.1SG axe.GEN.SG from you
'I will take the axe from you.' Russian (Koptjevskaja-Tamm & Wälchli 2001: 655)

The usage of genitive when the action is explicitly temporary ((28a) vs (28b)) is also noted for some North-Western Belarusian dialects, spoken in the area adjacent to the Lithuanian border:

(28) a. *pry-njas-i* *nož*
PVB-bring-IMP knife.ACC.SG
'Bring me the knife (implicitly: for a longer time).'
b. *pa-da-j* *naž-a*
PVB-give-IMP knife-GEN.SG
'Hand me the knife (just for a moment)' (Belarussian, BEL_1 and BEL_2)

Temporally restricted usage could also be illustrated by another type of clauses, where the genitive object refers to a specific purpose that is restricted in time. The meaning of temporal restriction is seen in *uchylić kapelusza.GEN* 'lift off ones hat', *dać buzi.GEN* 'give a kiss', *zapomnieć języka.GEN* 'forget one's tongue', *zapomnieć lekcję.GEN* 'forget a lesson' (Kempf 1970: 193). But the connection of genitive with a specific purpose can be seen in *dobyć miecza.GEN* 'draw a sword', where the sword is drawn with the aim of engaging in a fight. Holvoet cites the term *genetivus partitivus intentionalis*, originally coined by Marian Jurkowski, for a type of use referring to situations where the object is taken for the purpose of performing a well-defined, concrete action and illustrates this with an example for Polish dialect provided by Kempf (Holvoet 1991: 110):

(29) *Zlapie* *warzechy,* *wybije* *ci* *zęby.*
catch.FUT.1SG ladle.GEN.SG knock-out. FUT.1SG you tooth.ACC.PL
'I'll catch a ladle and knock out your teeth.' Polish (Kempf 1970: 1991)

Holvoet mentions that the meanings of different degrees of affectedness (which could be realized in slightly different ways as 'slight affectedness', 'temporal affectedness' or 'partial affectedness' expressed by partitive genitive) has the roots in Indo-European (Holvoet 1991: 111, Kempf 1970: 191). Different rules were applied for discrete objects and mass nouns; for discrete objects, it was probably a genuine partitive genitive, similar to that of Fennic. Later on, with the rise of the opposition between variable and constant quantification, the partitive was transformed into a genitive of quantity and the two meanings (genuine partitive genitive and genitive of quantity) became dissociated from each other. For discrete objects, the genitive could now denote a slight or superficial affectedness (for more details on the hypothesis of the historical development see Holvoet 1991: 111–112).

3.2 Surface-contact verbs

In a number of cases the use of the partitive or partitive genitive can be associated with a specific lexical class. An important difference is that between change-of-state and surface-impact verbs. Change-of-state verbs (such as English *break*) are verbs denoting a change from one state to another. Surface-contact verbs (like English *hit*) refer to physical contact between two objects, but from the use of these verbs it is not always obvious that the objects have undergone some essential change (Fillmore 1970: 130–131). In an abstract sense, surface-contact verbs identify *some* change as the person who was hit by someone is different from the person they were before the hitting occured.

A syntactic difference between change-of-state verbs and surface-contact verbs can be seen in English when the object is a body-part noun. The sentences with surface-contact verbs have paraphrases in which the possessor of the body part appears as the direct object and the body-part noun appears in a "locative prepositional phase" (Fillmore 1970: 131–132). Compare (30a) with the surface-contact verb to (30b) with the change-of-state verb:

(30) a. *I hit his leg.*
I hit him on the leg.

b. *I broke his leg.*
**I broke him on the leg.* (English, Fillmore 1970: 132)

Surface-contact verbs with partitive marking appear also in Baltic, Slavic and Fennic languages. Archaic Indo-European languages also have genitives:

(31) *Elábeto* *tês* *cheiròs* *autoû.*
take.AOR.MED.3SG ART.GEN.SG.F hand.GEN.SG 3.GEN.SG.M
'He took hold of his hand.' (Classical Greek, Goodwin 1898: 234)

In older modern Polish, some surface-contact verbs could also take genitive object:

(32) *zarzuciwszy* *wylotów* *i* *pogłaskawszy* *wąsa,*
throw_back.CVB mock_sleeve.GEN.PL and stroke.CVB moustache.GEN.SG
zaintonował […] *litanią*
intone.PST.3.SG.M litany.ACC.SG
'Having thrown back his mock sleeves and stroked his moustache, he intoned a litany.'
(Polish, Juliusz Słowacki, 19th c.)

The lexical meaning of the verb, rather than aspect, implies the slight degree of affectedness (Holvoet 1991: 109). An indirect trace of genitive with verbs of surface contact might also be seen in Russian:

(33) *kasnut-sa* *neba*
touch.INF-RFL sky.GEN.SG
'to touch the sky' (Russian, personal knowledge)

In modern Lithuanian, partitive genitive seems to be possible only with reflexive verbs as in (34a). Non-reflexive verbs would take accusative as in (34b). Normally the preference would be given for accusative (34b), but in some specific situations, when the person gets some impact, experience, knowledge about the nature of the object, partitive genitive would be used instead as in (34a):

(34) a. *pri-si-liesti* *dangaus*
PVB-RFL-touch.INF sky.GEN.SG
b. *pa-liesti* *dangų*
PVB-touch.INF sky.ACC.SG
'to touch the sky' (Lithuanian, personal knowledge)

Empirical data from older Lithuanian texts, e.g., from the 18th century, show clear traces of slight/partial affectedness expressed by a genitive object, as in (35):

(35) *<Bet priėjom wieną Sallą, Klaudą wadinnamą>,* *czonay* *wôs* *ne*
<But we have reached one island, which is called Klaud>, here scarely not
wôs *Waltiês* *gallėjome* *prigriebti.*
scarely boat.GEN.SG can.PST.1PL catch.INF
'<But we reached an island called Clauda>, here we could scarcely get hold of our boat.'
(Lithuanian, LT_18 corpus)

The sentence describes a situation in which a person could barely get hold of a boat, which was about to be carried away by water. The effort with which the object is seized is rendered by the use of the genitive marker for partial or superficial affectedness.

Finnish demonstrates much wider use of surface-contact verbs with partitive object, see (36a–c):

(36) a. *Jeesus kosketti hänen kättään.*
Jesus touch.PST.3SG he.GEN.SG hand.PAR.POSS.3SG
'Jesus touched his hand.' (Finnish, Raamattu, Matt 8, 15)

b. *Enkeli kosketti häntä <ja sanoi hänelle: "Nouse ja syö!">*
angel touch.PST.3SG he.PAR.SG <and said, "Get up and eat.">
'An angel touched him <and said, "Get up and eat.">
(Finnish, Raamattu, 1. Kun 5, 19)

c. *Hän taputti vanhan naisen selkää pienellä*
he pat.PST.3SG old.GEN.SG woman.GEN.SG back.PAR.SG small.ADESS.SG
kädellään ja sanoi pehmeästi <...>
hand.ADESS.POSS.3SG and say.PST.3SG kindly
'He patted the old woman's back with his small hand and said kindly <…>'
(Finnish[12])

Examples (36a–c) contain the surface-contact verbs *koskettaa* 'to touch' and *taputtaa* 'to pat'. There is some physical contact between two objects, marked with the partitive object. It is difficult to describe the nature of the change which the person undergoes when someone (e.g. an angel in (36b)) touches their hand.

Even though the use of the partitive object with Finnish surface impact verbs is a default, there are some exceptions; compare the difference between hitting something in (37a) (marked with the partitive) and hitting someone in such a way that the hitting causes death as in (37b), marked with the accusative:

(37) a. *Mooses kohotti sauvansa ja löi Niilin*
Mooses raise.PST.3SG staff.ACC.3POSS and strike.PST.3SG Nile.GEN.SG
vettä <…>
water.PAR.SG
'Moses raised his staff and struck the water of the Nile.'
(Finnish, Raamattu, 2. Moos 7: 20)

b. *Baesa löi hänet kuoliaaksi <...>*
Baasha beat.PST.3SG he.ACC.SG dead.TR.SG
'Baasha beat him to death.' (Finnish, Raamattu, 1. Kun 15, 27)

In the well-known example from Finnish involving shooting at someone and shooting someone dead (example (1) repeated here for the sake of convenience) different types of telic interpretation of the situation apply. The impact of the initial shooting intention is not clear. The verb *ampua* 'shoot' is a surface-impact verb whose meaning does not in itself imply a change. The opposition between two possible interpretations of the situation is marked with different object cases:

[12] https://tales.xperimentalhamid.com/fi/novel/the-proxy-bride-of-the-billionaire-chapter-531/

(1) a. *Ammuin karhun.*
shoot.PST.1SG bear.ACC.SG
'I shot the (a) bear.'
b. *Ammuin karhua.*
shoot.PST.1SG bear.PAR.SG
'I shot at the (a) bear (without killing it).' (Finnish, Kiparsky 1998: 2–3)

3.3 Scalar verbs

The culmination of the event, where the event reaches an endpoint, is the most important criterion for the choice between accusative and partitive for the Finnish direct object. This culmination is normally associated with telicity, but not every form of telicity entails culmination. In Finnish many verbs can show a distinction between *culminating* and *non-culminating telic behavior*. In most languages the non-culminating type would be represented by *telic scalar verbs*. This type is also known as a group of so called *degree achievement verbs*.[13] In English this type is represented by verbs like *widen, lengthen*. It was observed that these verbs have both telic and atelic properties: whilst atelic predicates are entailed by their progressive forms (Dowty 1979), some verbs in this group behave differently, e.g. the verb *lengthen* behaves like the atelic verbs (*Kim was lengthening the rope* entails *Kim has lengthened the rope*), whereas *straighten* behaves as telic in this respect (*Kim was straightening the rope* does not entail *Kim has straightened the rope*) (for more details see Hay et al. 1999: 127). The affected argument of telic scalar verbs undergoes a change in some property. In deadjectival verbs the change is in the property associated with the meaning of the adjectival base (Hay et al. 1999: 129).[14] The terminal point of the event can be identified with the following calculation: "the endpoint is that point at which the affected argument possesses a degree of the measured property that equals the initial degree to which it possessed this property plus the degree denoted by the difference value" (Hay et al. 1999: 133). When the difference value is not provided by overt linguistic material, it should be somehow inferred and boundedness is determined in other ways. Degrees are formalized as positive or negative intervals on a scale, where a scale is a set of points totally ordered along some dimension (Hay et al. 1999: 130–131), e.g. temperature, length, bad quality, strength etc.

Finnish verbs like *lämmittää* 'to warm up' are classified under telic scalar verbs (Larjavaara 2019: 229–231). The special feature of these verbs lie in their ability to have both partitive and accusative objects in sentences with discrete objects.[15] The verb *lämmittää* has two telic readings, one with the partitive (the non-culminational reading) and one with the accusative (the culminational reading). The use of the partitive in (38a) as opposed to the accusative in (38b) can be associated not only with an imperfective reading, but also with a non-culminational perfective reading:

---

[13] The term "degree achievement verbs" is taken from Dowty (1979) and has been criticized for inaccuracy as "degree achievements" show little evidence of being achievements at all (Hay et al. 1999: 143). Dowty claims that these verbs could be classified as achievements on certain semantic and syntactic grounds; Hay et al. argue that these verbs show the characteristics of accomplishments and activities (ibid).

[14] English adjectives fall into two classes: closed-range adjectives, which are associated with a scale with a maximal value, where maximality is relative to the adjective's polarity (e.g. straight, empty, dry) and open-range adjectives (e.g. long, bad, strong), for which it is not possible to identify maximal values on the scale (see Hay et al. 1999: 135–136 for a discussion about English adjectives). The telicity of degree achievements depends on the open-/closed range distinction. Degree achievements derived from open-range adjectives normally demonstrate atelic behavior.

[15] The use of Finnish telic scalar verbs is often dependent on the context or even on the dialectal background of the speaker. Sometimes direct object alternations between ACC vs. PAR with some certain telic scalar verbs could be seen as strange or even impossible. This serves as evidence that the group of telic scalar verbs is flexible and subjective interpretations of the events apply (Larjavaara 2019: 281).

(38) a. *Lämmitin* *saunaa.*
warm.PST.1SG sauna.PAR.SG
i. 'I warmed the sauna a bit.'
ii. 'I was warming up the sauna.'
b. *Lämmitin* *saunan.*
warm.PST.1SG sauna.ACC.SG
'I warmed up the sauna.' (Finnish, personal knowledge)

In (38a), there was a change from the initial state, but the change was not significant enough to reach the resultative end phase (Huumo 2013: 101). Telic scalar verbs usually allow a maximum possible effect, which is normally the optimal outcome of the event (Larjavaara 2019: 280–281). The progressive partitive as in (16a) and the irresultative partitive as in (38a) are similar in that the progressive partitive refers to an event that, if continued, finally reaches the endpoint (e.g. the book is read until the last page) and the same expectation could be linked with the irresultative partitive (the sauna can be warmed up to a point when it is warm enough). The irresultative partitive also indicates that the expected endpoint was never projected or never reached, because e.g. the action was interrupted by some outside event (for more details see Huumo 2013: 102).

The Finnish examples in (39), (40), (41) and (42) have the scalar structure of the adjectival base (*pitkittää* 'to lengthen', *pahentaa* 'to worsen', *vahvistaa* 'to strengthen', *lyhentää* 'shorten'). For a more detailed discussion of this type of verbs see Larjavaara (2019: 305–324):

(39) *<Ja mikä tulee olemaan loppuni, että vielä>* *pitkittäisin* *tämän*
< And what will be my end > prolong.COND.1SG this.GEN.SG
kaltaisen *sieluni* *elämää?*
alike.GEN.SG soul.POSS.1SG life.PAR.SG
'<And what will be my end> to further prolong the life of my soul like this?' (Finnish[16])

(40) *Jos* *yrität* *apuun,* *vain* *pahennat* *asiaa.*
if try.PRS.2SG help.ILL.SG only worsen.PRS.2SG case.PAR.SG
'If you try to help, you will just make the case worse.'
(Finnish, Raamattu, Sananl. 19, 19)

(41) *Nyt* *voit* *puhua,* *herrani,* *sinä* *olet* *vahvistanut*
now can.PRS.2SG talk.INF lord.POSS.1SG you have.PRS.2SG strengthen.PPA
minua.
me.PAR.SG
'Speek, my Lord, for you have strengthened me.' (Finnish, Raamattu, Dan. 10, 19)

(42) *Lyhensin* *hiuksiani.*
shorten.PST.1SG hair.PAR.PL.1POSS
'I shortened my hair.' (Finnish, personal knowledge)

In some cases, e.g., *pahentaa* 'worsen' the use of the accusative does not seem to be possible, probably because there is no absolute or normative degree of badness, which precludes the culminative use. The reason for the absence of an accusatival construction is obviously pragmatic in this case. For other verbs of this group alternations with accusative (representing the culminational reading) are possible, as in (43) and (44):

16 https://unski.blogaaja.fi/tuhlattu-aika/

(43) *Puheenjohtaja* *venytti* *puheensa* *kahden* *tunnin*
speaker.NOM strech.PST.3SG speech.ACC.POSS.3SG two.GEN.SG hour.GEN.SG
mittaiseksi.
long.TR.SG
'The chairman stretched his speech out over two hours.' (Finnish, p.c. K. Podshivalow)

(44) *Vahvistin* *aidan* *niin* *pitäväksi,* *etteivät* *villisiat*
strengthen.NOM fence.ACC.SG so firm.TR.SG that.NEG.3PL wild boar.NOM.PL
pääse *siitä* *läpi.*
pass.PRS.3SG this.ELAT.SG through
'I made the fence stronger so that the wild boars wouldn't get through it.'
(Finnish, p.c. K. Podshivalow)

The scalarity of verb meaning (and subjective expectations about the complete event) plays an important role for the morphosyntactic aspectual encoding (Tamm 2012: 19). In some cases, the exact endpoint cannot be verified by perception (for more examples and interpretations concerning endpoints see Larjavaara 2019: 217–230). Note, however, that the endpoints are categorized differently in subjective terms (which shows a clear link with pragmatics). Examples (45) and (46) are given to illustrate, that the exact endpoint is difficult to determine. In (45), the difference value of healthy and unhealthy lifestyle could be inferred (if generally accepted, that there is always a chance to make one's life healthier and healthier). In (46), with the verb *tahrata* 'to make something dirty' the exact endpoint of 'being dirty a bit' or 'being very dirty' is difficult to determine. Therefore in (46) only an abstract change is observed: when a person's hand becomes dirty because of iniquity, the person is not the same as before:

(45) <…> *he* *voivat* *muuttaa* *elämäntapojaan* *terveellisemmiksi*
they can.PRS.3PL change.INF lifestyle.PAR.PL.POSS.3PL healthy.COMP.TR.SG
'<…> they can change their lifestyles to healthier ones.' (Finnish[17])

(46) *Jos* *käsiäsi* *tahraa* *synti,* *heitä* *se*
if hand.PAR.POSS.2SG get-dirty.PRS.3SG iniquity.NOM throw.IMP it.ACC.SG
pois, <älä anna pahan asua majassasi. >
away, <…>
'If iniquity be in thine hand, put it far away, <and let not wickedness dwell in thy tabernacles>.' (Finnish, Raamattu, Job 11, 14)

In some situations the difference value is based on the context. For example, the length of the hair which I am shortening as in (42) might depend on some knowledge about hair during different periods of fashion (for more discussion on context-dependent telicity see Hay et al. 1999: 136–138).

Culminational telic and non-culminational telic uses in Finnish are further extended to other verbs that are not normally assigned to the class of telic scalar verbs, such as *avata* 'open'. An opposition between a culminational and a non-culminational reading is also observed here:

(47) a. *Hän* *avasi* *oven.*
3SG open.PST.3SG door.ACC.SG
'He opened the door.'
b. *Hän* *avasi* *ovea.*
3SG open.PST.3SG door.PAR.SG

[17] https://sansa.fi/kambodzalainen-nem-lin-haluaa-rakentaa-kirko/

'He opened the door for a while; he opened the door partly, set the door ajar); he was opening the door.' (Finnish, adapted from Kiparsky 1998: 8)

Example (47b), which is widely cited in literature (inter alia Larsson 1983: 87, Holvoet 1991: 109), can have progressive meaning (where the object is an incremental theme: 'he was opening the door'), but also several other meanings: 'he opened the door for a while', and also 'he partly opened the door' referred to as telic and perfective (as suggested by Kiparsky 1998: 8 in a similar example with opening the window; also Larjavaara 2019: 229). Example (47a) with the accusative object, is also characterized as telic, bounded and perfective and the semantic difference between these two sentences lies in identifying different endpoints. The telicity of these verbs in Finnish cannot be completely specified in terms of semantic or syntactic features and often derives from conventional implicatures:

(48) *Auto vaihtoi kaistaa.*
car.NOM change.PST.3SG lane.PAR.SG
'The car changed lanes.' (Finnish, Leino 1991: 171)

(49) *Kiristin ruuvia.*
tighten.PST.1SG screw.PAR.SG
i. 'I tightened the screw (a bit)
ii. 'I was tightening the screw.' (Finnish, Larjavaara 2019: 229)

In example (48) with verb *vaihtaa* 'change' the endpoint is based on other measurements (contrary to the example like with the verb *kirjoittaa* 'to write' (16a), where the writing event is linked with the last written sign of the letter being written) — changing the driving lane (but still staying on the road), tightening the screw to some extent, but not too much as in (49).

Estonian scholars also single out degree achievement verbs. Estonian transitive degree achievement verbs occur with the partitive object naturally, as it is the case with activity or accomplishment verbs (e.g. *build, paint, read* etc.); thus these verbs occur context-neutrally with partitive objects in durative sentences like (51) and primarily denote activities. Sentences (50)–(51) would qualify for an accomplishment and activity, sentence (52) illustrates an achievement-like reading (for more details on Estonian see Tamm 2012: 174–175):

(50) *Firma laiendas tee ühe tunniga.*
firm.NOM widen.PST.3SG road.ACC.SG one.GEN.SG hour.COM
'The firm widened the road in an hour.'

(51) *Firma laiendas teed kaks tundi.*
firm.NOM widen. PST.3SG road.PAR.SG two.NOM.SG hour.PAR.SG
'The firm was widening/widened the road for two hours.'

(52) *Firma laiendas teed ühe tunniga.*
firm.NOM widen. PST.3SG road.PAR.SG one.GEN.SG hour.COM
'The firm widened the road (a bit) in an hour.' (Estonian, Tamm 2012: 175–176)

In Russian dialects and Polish a few verbs can occasionally behave like the Finnish scalar telic verbs as well and take partitive genitive as object case. The examples are given for North Russian (53) and Polish (54):[18]

[18] In modern Polish most of such partitive genitives are now obsolescent or obsolete (Holvoet 1991: 107, Kempf 1970: 193). Kempf gives some examples from older Polish: *przytępić kosy.GEN* 'blunt a scythe', *przystrzyc*

(53) *Ja otvorju dverej.*
I open.FUT.1SG door.GEN.PL
'I will somewhat/partly open the door(s).' (Russian, Seržant 2020: 49)

(54) *Uchylił okna.*
open.PST.3SG window.GEN.SG
'He half-opened the window.' (Polish, Holvoet 1991: 107)

In the case of Polish *uchylić* the non-culminational telic meaning is lexicalized and case is also assigned lexically: genitive in older Polish and accusative in contemporary Polish. However, genitive case assignment is probably a trace of a former productive case alternation.

In Lithuanian, verbs showing non-culminational telic behavior with genitive partitive marking are hardly represented. There are no examples from modern Lithuanian, but it seems that such use was possible in older Lithuanian, as in (55):

(55) *<Nešęs velnias akmenį, didumo kaip gryčios, ir>* *sudaužyt*
<The devil was carrying a big stone, which was as big as the house, and> break.INF
norėjęs Anykščių bažnyčios.
want.PPA Anykščiai.GEN.PL church.GEN.SG
'<The devil was carrying a big stone, which was as big as the house, and> wanted to ravage the church.' (Lithuanian, corpus LT_19)

This old Lithuanian example, which is from the 19th century, could be interpreted in such a way that the church was subjected to partial destruction (the devil was carrying a big stone, but it affected only a part of the church, which was much bigger/stronger than a stone). However, this could also be a genitive of surface impact.

3.4 Conative verbs

The conative[19] alternation is a type of verb alternation between a verb construction indicating the completion of the action and a conative variant representing "an attempted action without specifying whether the action was actually carried out" (Levin 1993: 42, see also Goldberg 1995: 63). The notion of conative alternation is applied, in English, to certain semantic fields, e.g. verbs of contact by impact (*hit, kick*), see Levin (1993: 41):

(56) a. *John kicked the ball.*
b. *John kicked at the ball.* English (adapted from Levin 1993: 41)

Construction (56a) entails that the ball was hit while the corresponding conative construction (56b) does not imply that this aim was achieved. The conative construction marked with the preposition *at* signals that the event of kicking took place irrespective of the final result or success of the action (*James may have missed while trying to kick the ball*). In other words, the conative construction (56b) can be paraphrased as something like *James tried to kick the ball* (Levin 1993: 6). Conative alternations also convey different meanings in terms of intentionality (Anscombe 2000) and affectedness (Beavers 2006).

czupryny. GEN 'trim somebody's hair', where the action does not cover the whole object, but only parts of the object.

[19] The term *conative* comes from Latin *conor/conari* 'try, attempt'.

In Lithuanian, one could also find a few cases of lexical distinctions along the conativity dimension, like Lith. *įrodinėti* 'argue, try to prove' vs *įrodyti* 'prove', *įkalbinėti* 'try to persuade' vs *įkalbėti* 'persuade'. In Russian, conative meanings are coded, in some cases, with the alternation between imperfective/perfective verbs. In (57a) with the imperfective verb, the event of giving the money was not successful (the other person did not take the money), in (57b) with the perfective verb the final result is a success (the other person took the money):

(57) a. *Ja* *daval* *emu* *den'gi,* *<no on ne bral ih>.*
1SG give.PST.1SG he.DAT.SG money.ACC.PL, <…>
'I gave him money, <but he did not take it.>'
b. *Ja* *dal* *emu* *den'gi.*
1SG give.PST.1SG he.DAT.SG money.ACC.PL
'I gave him money.' (Russian, p.c. S. Podshivalow)

In Finnish conativity can be reflected in the form of the object. The conative alternation could be illustrated by examples (58a) and (58b). The alternation between partitive (58a) and accusative (58b) has nothing to do with partial affectedness: the person involved does not go 'a little bit' to a sauna, but either obeys the order given or not:

(58) a. *Käskin* *häntä* *saunaan.*
order.PST.1SG he.PAR.SG sauna.ILL.SG
'I ordered him to go to sauna (and he most probably went).'
b. *Käskin* *hänet* *saunaan.*
order.PST.1SG he.ACC.SG sauna.ILL.SG
'I ordered him to go to sauna (and he went).' (Finnish, Larjavaara 2019: 231)

Therefore, the difference between (58a) and (58b) is in the outcome of the event. In (58a) the emphasis is put on the action of giving the order to someone to go to a sauna (and the person most probably went to a sauna) while in (58b) the emphasis is both on the action and the outcome of the event (the person went to a sauna). For more explanations and examples see Larjavaara (2019: 231–232), where such verbs are classified under the group of telic fruition verbs (*teelis-suksessiiviset* in Finnish).

An analogous example involves the verb *suostutella* 'persuade', where the difference between the outcome of the action is also rendered by case marking:

(59) a. *Hän* *suostutteli* *ystävää* *elokuviin.*
3SG persuade.PST.3SG friend.PAR.SG movie.ILL.PL
'He tried to persuade a friend to go to the movies (but he did not go).'
b. *Hän* *suostutteli* *ystävän* *elokuviin.*
3SG persuade.PST.3SG friend.ACC.SG movie.ILL.PL
'He tried to persuade a friend to go to the movies (and he went).'
(Finnish, personal knowledge)

All the constructions discussed above describe a complex event involving at least two participants, where one is giving and another is (not necessarily) taking as in (57a–b), one is giving the order and another either obeying or not (as in (58a–b), one is trying to persuade another person to do smth. and the result is either successful or not (as in (59a–b). As a result, the irresultativity cannot be quantified, as in the case of degree achievements: in a long causal chain of successive sub-events, any of the necessary events can remain unrealized, leading to the irresultative character of the whole complex event.

4 Discussion

In the previous section a classification of the semantic factors that give rise to the variation of the case marking of discrete objects in Lithuanian and Finnish (and also other neighboring languages) were examined. The classification was based on verbs which normally assign accusative to discrete objects and the use of partitive or partitive genitive with such verbs is rather exceptional.

As Finnish does not have an overt aspect marking on the verb, partitive on the discrete object triggers the interpretation of an unbounded event (imperfective aspect or irresultativity). Lithuanian has overt marking of aspect on the verb (aspectual distinctions in Lithuanian are often expressed by the choice of verbal prefixes). Only plural and mass nouns can occur as objects, denoting indeterminate quantity, therefore the marking of discrete objects by partitive genitive in Lithuanian becomes problematic.

The question arises then whether the Finnish partitive case is used in situations/constructions where Lithuanian has an overt aspect marker on the verb? The resultative and irresultative readings of Finnish achievement verbs are marked with the accusative (= achievement) or partitive (= absence of the result) (see example (1a–b) repeated here for the sake of convenience):

(1) a. *Ammuin* *karhun.*
shoot.PST.1SG bear.ACC.SG
'I shot the (a) bear.'
b. *Ammuin* *karhua.*
shoot.PST.1SG bear.PAR.SG
'I shot at the (a) bear (without killing it).' (Finnish, Kiparsky 1998: 2–3)

In Lithuanian, the achievement or the absence of the result would be marked not on the object, but with different verbal prefixes as in (60a–b), which marks a difference in actionality:

(60) a. *Nu-šoviau* *lokį.*
PVB-shoot.PST.1SG bear.ACC.SG
'I shot the (a) bear.'
b. *Pa-šoviau* *lokį.*
PVB-shoot.PST.1SG bear.ACC.SG
'I shot at the (a) bear (without killing it).' (Lithuanian, personal knowledge)

Lithuanian verbal prefixes may offer a full range of possibilities to describe the event in a very detailed manner regarding the outcome of the result. To illustrate this, I give a non-prefixed Lithuanian verb *nešti* 'carry' with possible prefixes which modify the meaning of the verb and also the description of the result: *į-nešti* 'carry in', *iš-nešti* 'carry out', *per-nešti* 'carry along', *pri-nešti* 'carry at', *su-nešti* 'carry to', *už-nešti* 'carry up', *ap-nešti* 'carry around', *nu-nešti* 'carry to'. The opposition between *nešti* 'carry'*: į-nešti* 'carry in' is also the one of quantification. Apart from the description of the result of the event, prefixes may also reflect actional differences (i. e. differences in lexical aspect or Aktionsart), e.g., the prefix *pa-* in *pa-nešti* 'carry for some time' renders the verb perfective but atelic, and the boundedness associated with perfectivity is achieved through indication of an arbitrary boundary in time rather than through a change of state.

Historically, the partitive in Finno-Ugric was a spatial case with separative ('from') meaning (Kiparsky 1998: 32, Koptjevskaja-Tamm 2001: 534–535). Larjavaara (1991) shows that the object case variation in Finnish has developed in a logical manner, although there was some Baltic influence in the early stages. He argues that quantification is an older criterion than aspect and most probably the starting point from which the aspectual uses have developed. Whatever the factors involved in the historical development, the object marking of contemporary Finnish is based on

the culmination (or non-culmination) of the event: "the object of the sentence is total (= ACC) whenever and only when a positive sentence expresses a complete change of the event that has reached (or is reaching) its end-point. In all other cases (including transitive sentences denoting some extent of change, e.g. *lämmitin sauna*.PAR 'I heated the sauna (a bit)' or no change at all, e.g. *katsoin televisiota*.PAR 'I was watching TV'), the partial (= PAR) object is used" (Larjavaara 2019: 207). Therefore for Finnish the culmination of the event (which historically might have its roots in quantification) is the most important criterion for the assignment of the object case.

Could the Finnish partitive have taken over the same functions as the verbal aspectual marker in Lithuanian? The non-availability of irresultative meaning of the partitive genitive in Lithuanian and other Slavic languages might have something to do with the overt marking of aspect and Aktionsart on the verb. The irresultative reading of the object can often be marked in Baltic by a prefix reflecting an atelic Aktionsart.

This preliminary exploration of partitive genitive marking with discrete objects in Lithuanian shows inconsistent marking of irresultativity via case-marking. More diachronic research as well as research on the Lithuanian dialects would be needed to get a better picture of partitive marking of discrete objects in Lithuanian. However, the present research shows, that there is a clear difference between Finnish and Lithuanian for encoding irresultativity in discrete objects: Lithuanian strategy is to use aspectual prefixes, Finnish uses partitive marking.

5 Concluding remarks

In this article, the treatment of discrete objects in relation to object marking in Lithuanian and Finnish was investigated. As partitive genitive (or partitive) in object marking with discrete objects is also present in neighbouring languages (e.g. Polish *dać buzi*.GEN 'give a kiss', Karelian *antaa suuta*.PAR 'give a kiss', Russian *otvorit' dverej*.GEN 'partly open the door(s)'), one of the aims of this research was to investigate in which situations partitive or partitive genitive is interpreted as encoding an irresultative event in Lithuanian compared to other neighbouring languages, with a special focus on Lithuanian partitive genitive and Finnish partitive.

As was already shown in previous research, Lithuanian and Finnish have completely different criteria for assigning object cases (Laugalienė 2020). In Finnish the most important factor is the culmination of the event (or the absence of the endpoint), in Lithuanian quantification plays the most important role. In this article it was also shown that the endpoints in Finnish are of different types and could be categorized in subjective terms (i.e. the exact endpoint cannot be traced by perception, the result of the change in the mental state of the experiencer cannot be exactly verified, the event has temporal boundaries or it is not completed according to the judgment of the speaker etc.), which opens up an array of possibilities to use partitive marking for discrete objects very widely. As quantification plays the most important role for the object marking in Lithuanian, the possibilities to quantify discrete objects (*versus* mass nouns) are much more limited. In Lithuanian, partitive genitive can denote a slight affectedness (which could also be realized as affectedness limited in time). In modern Lithuanian, partitive genitive with discrete objects is obsolescent or obsolete, found only in some dialects. However, data from older Lithuanian show that it might have been more frequent than at present. In Finnish, variable quantification is associated with progressive and imperfective readings. Quantification undoubtedly plays an important role in the Finnish aspect both from the diachronic and synchronic point of views (Larjavaara 2019: 209).

Finnish does not have overt aspect marking on the verb, whereas Lithuanian has markers on verbs (mainly prefixes). Hence, Finnish partitive on the discrete object triggers the interpretation of an unbounded event (imperfective aspect or irresultativity). As Lithuanian has overt marking of aspect and actionality (lexical aspect, Aktionsart) on the verb, marking of irresultativity in the case form of the object often becomes redundant. Therefore the non-availability of irresultative

meaning of the partitive genitive in Lithuanian (and other Slavic languages) might partly be due to overt marking of aspect and actionality on the verb.

To conclude, Finnish and Lithuanian both have a way of morphologically marking partially affected discrete objects but do so in different ways. Lithuanian encodes aspect on the verb and so marks partial affectedness that way, but Finnish uses partitive instead. The hypothesis at the beginning of this research was, that Finnish and Lithuanian might use different strategies for encoding irresultativity in discrete objects (Lithuanian has possibilities to express irresultativity with aspectual prefixes, Finnish expresses irresultativity with partitive marking). The results of the research confirm this hypothesis: no consistent marking of irresultativity via case-marking would be expected in Lithuanian, because that would be redundant. The fact that there are only 8 examples in Table 1 (occurrences of verbs with partitive genitive for discrete objects in Lithuanian corpora, which contains both examples from old Lithuanian and Lithuanian dialects) demonstrates that the partitive genitive strategy for discrete objects never completely developed in Lithuanian.

Sources

BEL_1 = Avanesaŭ, R. I., K. K. Krapiva & Ju. F. Matskevich (eds.). 1963. *Dyjalektalahichny atlas belaruskaj movy* [A dialect atlas of Belarusian language]. Vol. 3. Minsk: Vydavetstva Akademii Navuk BSSR, map 211.

BEL_2 = Avanesaŭ, R. I., K. K. Krapiva & Ju. F. Matskevich (eds.). 1963. *Dyjalektalahichny atlas belaruskaj movy. Ustupnyja artykuly, davedchyja materyjaly i kamentaryi da kart* [A dialect atlas of Belarusian language. Introductory chapters, reference materials, and commentary on maps]. Minsk: Vydavetstva Akademii Navuk BSSR, 756-757.

LT_16 = 1547, Mazvydas Martinas, Catechismus; 1573, Wolfenbüttelio postilė, B; 1575, Vilentas Baltramiejus, Enchiridion z; 1585, Jonas Bretkūnas, Biblija, Patarlių knyga (Pat); 1586, Jonas Bretkūnas, Biblija, Mokytojo knyga (Mok); 1589, Jonas Bretkūnas, Giesmės Duchaunos; 1589, Jonas Bretkūnas, Kancionalas; 1589, Jonas Bretkūnas, Kolektos; 1590, Jonas Bretkūnas, Biblija, Antra metraščių (kronikų) knyga (2 Met); 1590, Jonas Bretkūnas, Biblija, Danieliaus knyga (Dan); 1590, Jonas Bretkūnas, Biblija, Giesmių giesmės knyga (Gg); 1590, Jonas Bretkūnas, Biblija, Izaijo knyga (Iz); 1590, Jonas Bretkūnas, Biblija, Jeremijo knyga (Jer); 1590, Jonas Bretkūnas, Biblija, Jobo knyga (Job); 1590, Jonas Bretkūnas, Biblija, Pirma metraščių (kronikų) knyga (1 Met); 1590, Jonas Bretkūnas, Biblija, Raudų knyga (Rd); 1585, Jonas Bretkūnas, Biblija, Josua; 1585, Jonas Bretkūnas, Biblija, Judges; 1585, Jonas Bretkūnas, Biblija, Kings 1; 1585, Jonas Bretkūnas, Biblija, Kings 2; 1585, Jonas Bretkūnas, Biblija, Ruth; 1585, Jonas Bretkūnas, Biblija, Samuel 1; 1585, Jonas Bretkūnas, Biblija, Samuel 2; 1591, Jonas Brertkunas Altes Testament; 1591, Jonas Bretkunas, Postilla 1; 1591, Jonas Bretkunas, Postilla 2; 1595, Mikalojus Daukša, Katekizmas; 1599, Mikalojus Daukša, Postilė.

LT_18 = 1700, Martinas Lutheris, Mažasis katekizmas; 1701, Samuelis Bitneris, Naujasis Testamentas; 1704, Danielius Kleinas, Frydrichas Šusteris, Naujos giesmių knygos, Naujos maldų knygelės; 1704, Danielius Kleinas, Frydrichas Šusteris, Naujos giesmių knygos; 1704, Danielius Kleinas, Frydrichas Šusteris, Naujos maldų knygelės; 1705, Jonas Jaknavičius, Ewangelie Polskie y Litewskie; 1706 Jokūbas Perkūnas, Wohlgegründetes Bedenken über die ins Litauische übersetzten zehn Fabeln Aesopi und derselben passionierte Zuschrift; 1706, Jonas Šulcas, Ezopo pasakėčios; 1706, Michaelis Mörlinas, Principium primarium in lingua Lithuanica; 1709, Martinas Lutheris, Mažasis katekizmas; 1716–1744, Jokūbas Brodovskis, Litauische Sprichwörter und Rätsel; 1727, Ruigis, Naujasis Testamentas; 1728, Psalteras Dovydo; 1729, Johannas Anastasijus Freylinghausenas, Davadnas mokslas apie dūšios išganymą; 1750, Jonas Jaknavičius, Ewangelie Polskie y Litewskie; 1775, Gotfridas Ostermejeris, Agenda.

LT_19 = 1869, Motiejus Valančius; 1897, Pranas Vaičaitis; 1823, Silvestras Valiūnas; 1807, Johannas Arndtas, Šešios knygos apie tikrą krikščionumą (pirma knyga); 1815, Dionizas Poška, Mužikas Žemaičių ir Lietuvos; 1816, Juozapas Arnulfas Giedraitis, Naujas įstatymas; 1823, Simonas Stanevičius, Historija šventa; 1841, Keturios gražios giesmės; 1841, Ketury naujo giesmė; 1841, Penkis naujo giesmė; 1845, Eikš prie Jėzaus; 1859, Baranauskas, Anykščių šilelis; 1896, Antanas Baranauskas, Kalbomokslis lietuviškos kalbos

LT_20 = 1896-1933, Vaižgantas; 1919, Basanavičius; 1913, Ignas Šeinius_Raštai_9; 1903, Lazdynų pelėda; 1933, Ignas Šeinius_Raštai_10; 1937, Ignas Šeinius, Raštai_5; 1942, Ignas Šeinius_Raštai_6

Raamattu Dan. = Danielin kirja (Raamattu 1992, WSOY, Porvoo, Helsinki, Juva)

Raamattu Jer = Jeremian kirja (Raamattu 1992, WSOY, Porvoo, Helsinki, Juva)

Raamattu Job = Jobin kirja (Raamattu 1992, WSOY, Porvoo, Helsinki, Juva)

Raamattu Matt. = Evankeliumi Matteuksen mukaan (Raamattu 1992, WSOY, Porvoo, Helsinki, Juva)

Raamattu, 2. Moos 7: 20 = Toinen Mooseksen kirja (Raamattu 1992, WSOY, Porvoo, Helsinki, Juva)

Raamattu Sananl. = Sananlaskujen kirja (Raamattu 1992, WSOY, Porvoo, Helsinki, Juva)

Raamattu 1. Kun = Ensimmäinen kuninkaiden kirja (Raamattu 1992, WSOY, Porvoo, Helsinki, Juva)

URB 2013 = Buteikytė-Urbanavičienė, Vanda. 2013. *Pro Usioginę vieškeliu, tolin...* Ukmergė: Valdo leidykla.

References

Ambrazas, Vytautas, Dumašiūtė, Zelma, Juodelytė D., Kniūkšta, Pranas, Labutis, Vitas. 1976. *Lietuvių kalbos gramatika. III tomas. Sintaksė* [Lithuanian grammar. 3rd Volume. Syntax]. Vilnius: Mokslas.

Ambrazas, Vytautas (ed.). 1997. *Lithuanian Grammar.* Vilnius: Baltos Lankos.

Anscombe, Gerthrud E. M. 2000. *Intention.* Cambridge: Harvard University Press.

Beavers, John. 2006. *Argument/Oblique Alternations and the Structure of Lexical Meaning.* Doctoral Dissertation, Stanford University.

Blake, Barry J. 2004. *Case.* 2nd edition, Cambridge: Cambridge University Press.

Buslaev, Fedor I. 1959. *Istoričeskaja grammatika russkogo jazyka* [Historical grammar of Russian language]. Moskva, Učpedgiz.

Dowty, David. 1979. *Word meaning in Montague Grammar.* Dordrecht: Reidel.

Fillmore, Charles J. 1970. The grammar of hitting and breaking. In R. Jacobs & P. A. Rosenbaum (eds.), *Readings in English transformational grammar.* 120–133. Waltham, MA: Ginn.

Goldberg, Adele E. 1995. *Constructions: a construction grammar approach to argument structure.* Chicago [Ill.] – London: The University of Chicago Press.

Goodwin, William W. 1898. *Greek Grammar.* Boston: Ginn&Company.

Hay, J., Kennedy, C., & Levin, B. 1999. Scalar Structure Underlies Telicity in "Degree Achievements". In T. Mathews and D. Strolovitch (eds.), *SALT IX, CLC Publications*, 127–144. Ithaca.

Heinämäki, Orvokki. 1984. Aspect in Finnish. In Casper de Groot and Hannu Tommola (eds.), *Aspect bound: A voyage into the realm of Germanic, Slavonic and Finno-Ugrian aspectology*, 153–177. Dordrecht: Foris Publications.

Holvoet, Axel. 1991. *Transitivity and clause structure in Polish: a study in case marking.* Warszawa: SOW.

Huumo, Tuomas. 2013. On the many faces of incompleteness: Hide-and-seek with the Finnish partitive object. *Folia Linguistica* 47(1). 89–111.

ISK 2004 = Hakulinen, Auli, Maria Vilkuna, Riitta Korhonen, Vesa Koivisto, Tarja Riitta Heinonen, and Irja Alho. 2004. *Iso suomen kielioppi* [Descriptive Grammar of Finnish]. Suomalaisen Kirjallisuuden Seuran toimituksia 950. Helsinki: Suomalaisen Kirjallisuuden Seura.

Jablonskis, Jonas & Baublys, Rostislavas. 1957. *Rinktiniai raštai* [Selected articles]. Vilnius: Valstybinė politinės ir mokslinės literatūros leidykla [State Publiching House of Political and Scientific Literature].

Kempf, Zdzisław. 1970. Rozwój i zanik polskiego partitiwu [The development and the dissapearance of the Polish partitive]. In Jan Safarewicz (ed.), *Język Polski: organ Towarzystwa Miłośników Języka Polskiego* 3 [Polish Language: the Society of Polich Language Enthusiasts], 181–194. Towarzystwo Miłośników Języka Polskiego [The Society of Polich Language Enthusiasts].

Kiparsky, Paul. 1998. Partitive Case and Aspect. In Miriam Butt & Wilhelm Geuder (eds.), *The projection of arguments: lexical and compositional factors*, 265–307. Stanford: CSLI Publications.

Koptjevskaja-Tamm, Maria. 2001. "A piece of the cake" and "A cup of tea": partitive and pseudo-partitive nominal constructions in the Circum-Baltic languages. In Östen Dahl & Maria Kopjevskaja-Tamm (eds.), *The Circum-Baltic Languages: Typology and Contact*, Vol 2, 523–568. Amsterdam: Benjamins.

Koptjevskaja-Tamm, Maria & Bernhard Wälchli. 2001. The Circum-Baltic languages: An areal-typological approach. In Östen Dahl and Maria Koptjevskaja-Tamm (eds.), *The Circum-Baltic Languages: Typology and Contact*, Vol. 2, 615–750. Amsterdam, Philadelphia: John Benjamins.

Kuryłowicz, Jerzy. 1971. Słowiański genetivus po negacji [Slavic genitive in negation]. In *Sesja Naukowa Międzynarodowej Komisji Budowy Gramatycznej Języków Słowiańskich* [Scientific Session of the International Commission for the Structure of the Grammar of Slavi Languages] (= Polska Akademia Nauk. Oddz. w Krakowie, Prace Komisji Słowianoznawstwa 23), 11–15., Wrocław: Ossolineum.

Larjavaara, Matti. 1991. Aspektuaalisen objektin synty [The Origin of the Aspectual Object]. *Virittäjä* 95. 372–408.

Larjavaara, Matti. 2019. *Partitiivin valinta* [The Choice of the Partitive]. Helsinki: Suomalaisen kirjallisuuden seura.

Larsson, Lars-Gunnar. 1983. Studien zum Partitivgebrauch in den ostseefinnischen Sprachen [Studies of the use of the partitive in Balto-Finnic languages]. Acta Universitatis Upsaliensis. *Studia Uralica et Altaica Upsaliensia* 15. Uppsala: Uppsala University.

Laugalienė, Asta. 2020. Partitivity and object marking in Finnish and Lithuanian. *Philologia Estonica* 5. 236–267. https://doi.org/10.22601/PET.2020.05.08

Leino, Pentti. 1991. *Lauseet ja tilanteet. Suomen objektin ongelmia* [Sentences and situation. The problems of the Finnish object]. Helsinki: Suomalaisen Kirjallisuuden Seura.

Levin, Beth. 1993. *English Verb Class and Alternations: A Preliminary Investigation.* Chicago: The U of Chicago P.

Luraghi, Silvia & Kittilä, Seppo. 2014. Typology and diachrony of partitive case markers. In Silvia Luraghi & Tuomas Huumo (eds.), *Partitive Cases and Related Categories (*Empirical Approaches to Language Typology 54), 17–62. Berlin & Boston: DeGruyter.

Miestamo, Matti. 2014. Partitives and negation: a cross-linguistic survey. In Silvia Luraghi & Tuomas Huumo (eds.), *Partitive Cases and Related Categories* (Empirical Approaches to Language Typology 54), 63–86. Berlin & Boston: DeGruyter.

Seržant, Ilja A. 2014. The Independent Partitive Genitive in Lithuanian. In Axel Holvoet & Nicole Nau (eds.), *Grammatical Relations and their Non-Canonical Encoding in Baltic (Valency, Argument Realization and Grammatical Relations in Baltic 1)*, 257–299. Amsterdam & Philadelphia: John Benjamins.

Seržant, Ilja A. 2015. The Independent Partitive as an Eastern Circum-Baltic isogloss. *Journal of Language Contact* 8. 341–418.

Seržant, Ilja. 2020. Typology of partitives. *Journal of Linguistics* 59, 4, 5–67.
Tamm, Anne. 2012. Scalar verb classes. Scalarity, Thematic Roles and Arguments in the Estonian Aspectual Lexicon. *Biblioteca di studi di filologia moderna* 14, Firenze University Press.
Vainikka, Anne & Maling, Joan. 1996. Is partitive case inherent or structural? In Jacob Hoeksema (ed.), *Partitives: Studies on the Syntax and Semantics of Partitive and Related Constructions,* 179–208. Berlin: MoutonDe Gruyter.

Asta Laugalienė
Vilnius University, Vilnius
asta.laugaliene@flf.vu.lt

Computation and the Justification of Grammars

Pauli Brattico

This methodological note revisits the original criteria proposed by Chomsky (1965) for the justification of grammars and suggests that modern computational methods could provide a useful tool for such purposes. Fully rigorous methods can help assessing observational, descriptive, explanatory and psycholinguistic adequacy of formally rigorous linguistic theories. The methodology is applied to the study of Finnish agreement.

Keywords: *agreement, computational linguistics, Finnish, linguistic theory, justification*

1 Introduction

It is common knowledge that the less advanced sciences, such as linguistics or sociology, do not generate the kind of cumulative knowledge characteristic of the more advanced disciplines. It is possible to accomplish productive and distinguished careers in the "soft sciences" while maintaining opposing views on virtually every issue, no matter how fundamental. In linguistics, for example, no agreement exits on questions such as what language is.[1]

One possible explanation for this discrepancy is that the more advanced sciences use a different method for justification than the less advanced ones. In the former, theories are connected to observations by deductive calculation. This system, first used in its present form by Galileo and later institutionalized by Newton, possesses an unrivaled epistemological power because it removes opinion from scientific justification. The medieval criteria of thought experiment, common sense, human intuition, authority, popularity, institutional structure, author reputation, political correctness, sociology of science, imagination, or any type of Augustine's "divine illumination" play no role in justification in these fields (although they do play a role in other affairs such as discovery). Thus, in linguistics, too, we should aim to "construct a formalized general theory of linguistic structure" because by "pushing a precise but inadequate formulation to an unacceptable conclusion, we can often expose the exact source of this inadequacy and, consequently, gain a deeper understanding of the linguistic data. More positively, a formalized theory may automatically provide solutions for many problems other than those for which it was explicitly designed" (Chomsky 1957: 5). In fact, a grammar that is "perfectly explicit" and does not rely "on the intelligence of the understanding reader" (Chomsky 1965: 4) is considered to be "generative," hence the term "generative grammar"

[1] It is possible to accomplish distinguished professional careers in linguistics by believing that language is not a natural phenomenon at all or that it is a biological property of the human brain; that it is based on innate properties of the human cognitive architecture or that it is learned by simplest Pavlovian association; that it has autonomous syntax or that its principles are covered by general cognitive or even pragmatic principles; that it has recursive syntax or only nonrecursive components; that is has no qualitative differences when compared to nonhuman 'languages' or that the human language is biologically unique; that there is a specialized language faculty or none exists, and so on. Virtually any imaginable position can be and has been entertained despite the fact that the data remains the same.

refers to a theory that is formal in this exact sense. Yet, such methods are almost never employed, perhaps due to the complexity of the required linguistic calculations making them unfeasible from the point of view of practical research projects. I propose in this note that modern computational tools provide a feasible way out of this methodological difficulty.

2 Background

First we must define the term "computational linguistic theory" to dispel some myths. A computational linguistic theory must satisfy two conditions: it must be (i) unambiguous and (ii) expressed in a machine-readable way. The requirement that a scientific theory must be unambiguous means that it does not rely on notions or assumptions that are open to interpretation. This allows the researcher to connect the theory with observation in a way that does not leave room for opinion, disagreement or logical gaps. In addition, when scientists put forward ambiguous theories they must be implicitly or explicitly assuming that an unambiguous formulation exists; believing to the contrary would be tantamount to saying that the theory must have some poetic quality making it necessarily ambiguous. Condition (i) is therefore nonnegotiable. Condition (ii) imposes an additional requirement: the theory must be provided in a machine-readable format or at the very least implemented in such notation. This allows the researcher to use a computer to test the theory against observation by using deductive calculation. In short, a computational linguistic theory is an ordinary linguistic theory formulated in some unambiguous notation that a machine can understand. No other properties are at stake.

While I will claim that linguistics can benefit from the use of rigorous computational methods in justification, this is also the only thing I want to claim. I do not propose to eliminate human intuition from the scientific discovery process or from any subject matter consideration, or to replace the 17th century scientific method with 19th century positivism that suspended all abstract theorizing. My concern is justification: how to bridge the theory, discovered by whatever mystical process, with observations, acquired by some means I do not want to restrict. Similarly, the point is not to replace linguistic theorizing with data mining or apply computational techniques to datasets in the hopes of discovering surface correlations. While computational discovery procedures can be useful in some contexts, they are irrelevant to the matters discussed in this article. Finally, the medieval method that relies on divine illumination or some other form of superior human cognitive capacity in justifying scientific hypotheses is, whatever faults it has, able to generate true theories. It is also able to produce interesting observations. One can discover groundbreaking truths even by pure luck. What the medieval method is unable to produce is agreement.

To show that the proposed computational methodology is feasible within the context of a real linguistic research project, I will use a concrete computational framework in this article as an example. The example system is a Python-based program that I wrote to provide an idealized "brain model" of a speaker of any language allowing the researcher to embed it with linguistic analyses and to test them by calculation. The framework consists of several interconnected components, the most important shown in Figure 1.

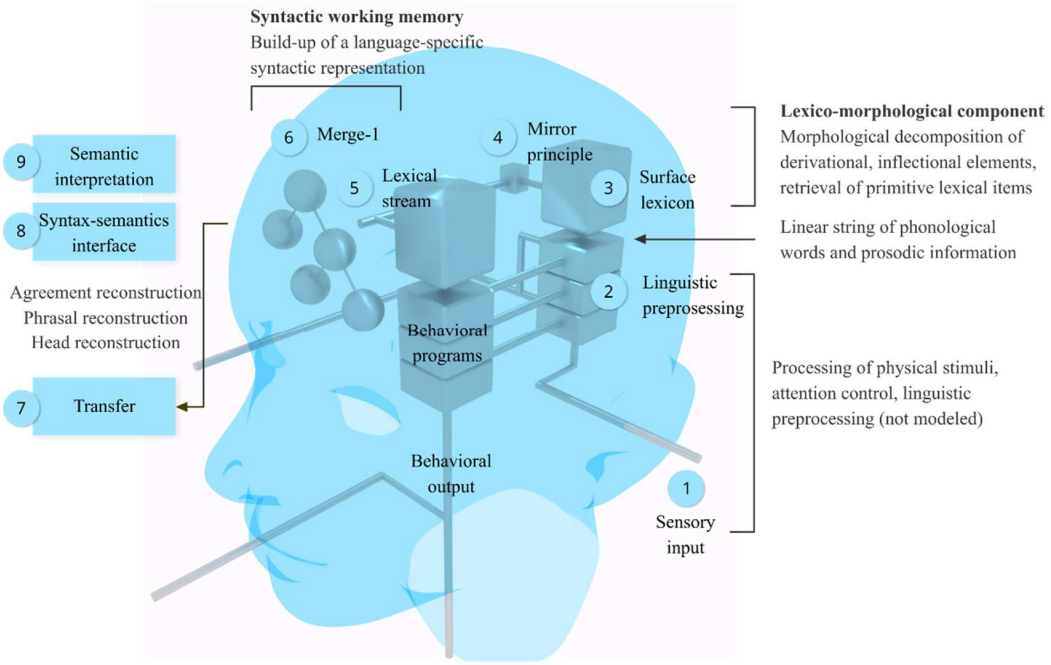


Figure 1: *Diagram of the various components of the Python-based brain model used as an example in this article. Input sentences are linear sequences of phonological words (1–2), which are processed by a lexico-morphological component performing lexical retrieval and morphological decomposition (3), followed by mapping of lexical items into syntax (4–5), generation of parsing solutions (6) and transfer (7–8) into systems of semantic interpretation (9)*

The model reads input sentences from left to right, retrieves each input word from the lexicon (3, Figure 1), merges them into a partial phrase structure in the current active working memory (6), and, once all words have been consumed from the input, transfers the calculated result to the syntax-semantics interface (7, 8) for evaluation and semantic interpretation (9).[2] It therefore maps phonological input sentences from the sensory interface(s) into sets of syntactic analyses and semantic interpretations. Input sentences that are judged ungrammatical are not interpreted semantically and are marked ungrammatical in the calculated output. The architecture was developed on the basis of earlier work by Phillips (1996) and is documented in Brattico (2019a).

Consider an input sentence *the horse raced past the barn*. The sentence is consumed one word at a time from left to right (1a) while each lexical item, retrieved on the basis of the phonological word in the input, is merged incrementally to a partial syntactic representation that exists in the algorithm's working memory (1b–c).

(1)	a.	*the*	*	*horse*	*	*raced*	*	*past*	*	*the*	*	*barn*	(Input)
	b.	↓		↓		↓		↓		↓		↓	(Merge)
	c.	[[the		horse]		[raced[		past		[the		barn]]]]	(Result)

[2] It is usually assumed in the cognitive sciences that the human conceptual system is not language-specific. This presupposes that the syntactic processing pathway eliminates language-specific features from the input before interpretation. Transfer (7, Figure 1) accomplishes this task. It incorporates a reverse-engineered chain creation algorithm, thus much of the linguistic theorizing currently in focus in generative theorizing is encapsulated inside this component.

If no more words appear, (1c) is interpreted semantically. Suppose however that there is one more word, an intransitive verb *fell*, in the input. Because there is no grammatically legitimate position for the intransitive verb in (1c), the algorithm reconsiders earlier parsing decisions by backtracking. Ambiguities are discovered in the same way. Backtracking finds all possible grammatical analyses and semantic interpretations for any given input string that are consistent with the linguistic hypotheses incorporated into the model. Ungrammatical input sentences are marked in the output as such. We examine this process in Section 4.

Whether this particular model is correct or even plausible is not important. Most linguists would judge its underlying assumptions as misguided. An expert who reviewed a manuscript advocating the above model criticized it as completely clueless, thus "nothing whatsoever" justified it, in his or her opinion. There was no "conceptual realm […] in which it might make some sense or have some application", and furthermore there was no explanation, according to this expert, of "who or what is supposed to carry out these operations". The idea that language is involved with neurocognitive computations of some sort was considered so outrageous as to be incomprehensible. In that reviewer's expert opinion, then, in order to proceed "we need to be looking at areas of linguistic inquiry […] very far removed from anything this author is interested in". This complete lack of agreement on every aspect of every theory aside (a standard feature of the less advanced sciences), my point is not to argue whether the model is correct or incorrect; rather, the point is that it is not justified by the type of subjective opinion exhibited by the reviewer in these remarks. Let us examine how it is justified.

3 Justification of grammars

The model described in Section 2 maps input sentences into sets of semantic interpretations and, if no mapping is found, judges them ungrammatical. It therefore captures a notion of linguistic competence by partitioning any set of input sentences into grammatical and ungrammatical, and by providing the former with a grammatical analysis (or several analyses). The Python implementation makes this process automatic.

Let us consider the word order study reported by Brattico (2020). Twenty Finnish seed sentences were selected that represent basic construction types in Finnish. All possible word order permutations were mechanically generated from the seed sentences. This method generated 119800 unique word orders. We can regard this set as a minimal word order corpus for this language, containing all possible word order permutations derived from a set of relatively simple seed constructions. Working with a corpus of this size with the traditional paper-and-pencil methodology would be infeasible, but it presents a trivial task for the computer. The recognition algorithm enriched with grammatical word order principles used less than a day to calculate structural analyses and grammaticality judgements for each sentence in this corpus. The model was justified, to the extent that it was, by matching the calculated output with native speaker judgments.

Let us break the process into several stages. I build an ad hoc test corpus of 2038 construction types that represent a wide variety of linguistic constructions in Finnish, English and Italian. The test corpus, therefore, contains data considered relevant to the subject matter under study. If we were concerned with word order, then all logically possible word order permutations from some set of basic constructions should appear in this file. If the focus was on some specific phenomenon such as *pro*-drop, we should use finite clauses exemplifying possible and impossible *pro*-drop configurations. If the study

aims for establishing crosslinguistic generalizations, then we include sentences from several languages. The contents of the test corpus used for the purposes of the present article are listed in Table 1.

GROUP	DESCRIPTION
1	**Basic construction types**
1.1	Basic verb classes
1.1.1	Intransitive verbs
1.1.2	Transitive verbs
1.1.3	Ditransitive verbs
1.1.4	Ditransitive verbs plus PP argument
1.1.5	Two PP arguments
1.2	Special finite elements
1.2.1	Auxiliary-like negation
1.2.2	Modal constructions, Neg + Modal + V
1.2.3	Pure tensed auxiliaries
1.3	Clausal infinitivals
1.3.1	Clausal infinitivals in English
1.3.2	English OC-constructions
1.3.3	Finnish clausal infinitivals (A/inf, VA/inf)
1.4	Nominals
1.4.1	Basic DP constructions
1.4.2	N + clausal infinitival
1.5	Adpositions
1.5.1	Prepositions, postpositions
1.6	Embedded *that*-clauses
1.6.1	Embedded *that*-clauses in Finnish and English
1.7	Relative clauses
1.7.1	Subject relativization
1.7.2.	Object relativization
1.8	Lexical ambiguity
1.8.1	Lexical ambiguity tests (frequency based)
2	**Adjuncts and adjunction structures**
2.1	PP adjuncts
2.1.1	Postverbal PP adjunct constructions
2.2	Adjectives
2.2.1	DP-internal adjectives (Finnish, English)
2.3	Clausal adjunct infinitivals in Finnish
2.3.1	MA-infinitivals
2.3.2	ESSA-infinitival
2.3.3	TUA-infinitival
2.3.4	E-infinitival
3	**A-bar (operator) movement and pied-piping**
3.1	Basic interrogatives
3.1.1	Subject and object interrogatives
3.2	Pied-piping
3.2.1	Pied-piping in Finnish and English
3.3	Islands
3.3.1	*CED-effects
3.3.2	*Extraction from embedded wh-clause
3.3.3	*Extraction from DP
3.3.4	*Extraction from embedded subject position
3.4	Left-peripheral C-features (Finnish only)
3.4.1	All agglutinative combinations
3.4.2	Single C-feature (subjects, objects)

9.6.1	Intransitives
9.6.2	Transitives
9.6.3	Ditransitives
9.6.4	Neg-to-C
9.6.5	LHM
10	**Clitics (Italian)**
10.1	Direct object clitics
10.2	Two-verb constructions
10.3	Three-verb constructions
10.4	Clitic agreement constructions and tests
10.5	Indirect clitic arguments
10.6	Clitic clusters
10.7	Restructuring
10.8	Reflexives

Table 1: *Test sentences (2038 in total). A category that exemplifies only ungrammatical sentences is marked with an asterisk*

One strength of this framework is that all hypotheses and theories aspiring to explain some linguistic phenomena can potentially agree to a common dataset, as defined by the test corpus. Another benefit is that everybody will come to the arena with the same requirement: propose a formula that calculates the same data. We will also eliminate a situation where two linguistic theories compete against each other while working, implicitly or explicitly, with different datasets.

Next, a script was deployed that read all sentences from the test corpus and fed them to the idealized speaker model (Figure 1), which then processed the sentences on the basis of the linguistic principles hypothesized by the author. In this way, we can examine if the hypothesis replicates the grammaticality judgments of human informants and "presents the observed data correctly" (Chomsky 1964: 28). This constitutes a minimal criterion for any scientific hypothesis, in any field. To do this, we create a gold standard and compare it with the calculated output. An example comparison, when I ran the test corpus through an algorithm that existed at the time of this writing, is provided in Figure 2. The gold standard is on the left, model output is on the right. The rightmost yellow column shows the comparison over the whole test corpus. Discrepancies are highlighted in red.

If a sentence is judged ungrammatical, then no output apart from the judgment itself is produced. Ungrammatical sentences have neither well-defined phrase structure representations nor semantic interpretations. To find out why some sentence was judged ungrammatical, we consult a derivational log file that stores all linguistically relevant computational steps executed during the calculations. Let us examine the expression *se talo* 'that.NOM house.NOM' (#197, Figure 2, line 312) that the model judged wrongly as ungrammatical. We locate the input from the derivational log file and examine what happened when the model processed that input. This is shown in Figure 3.

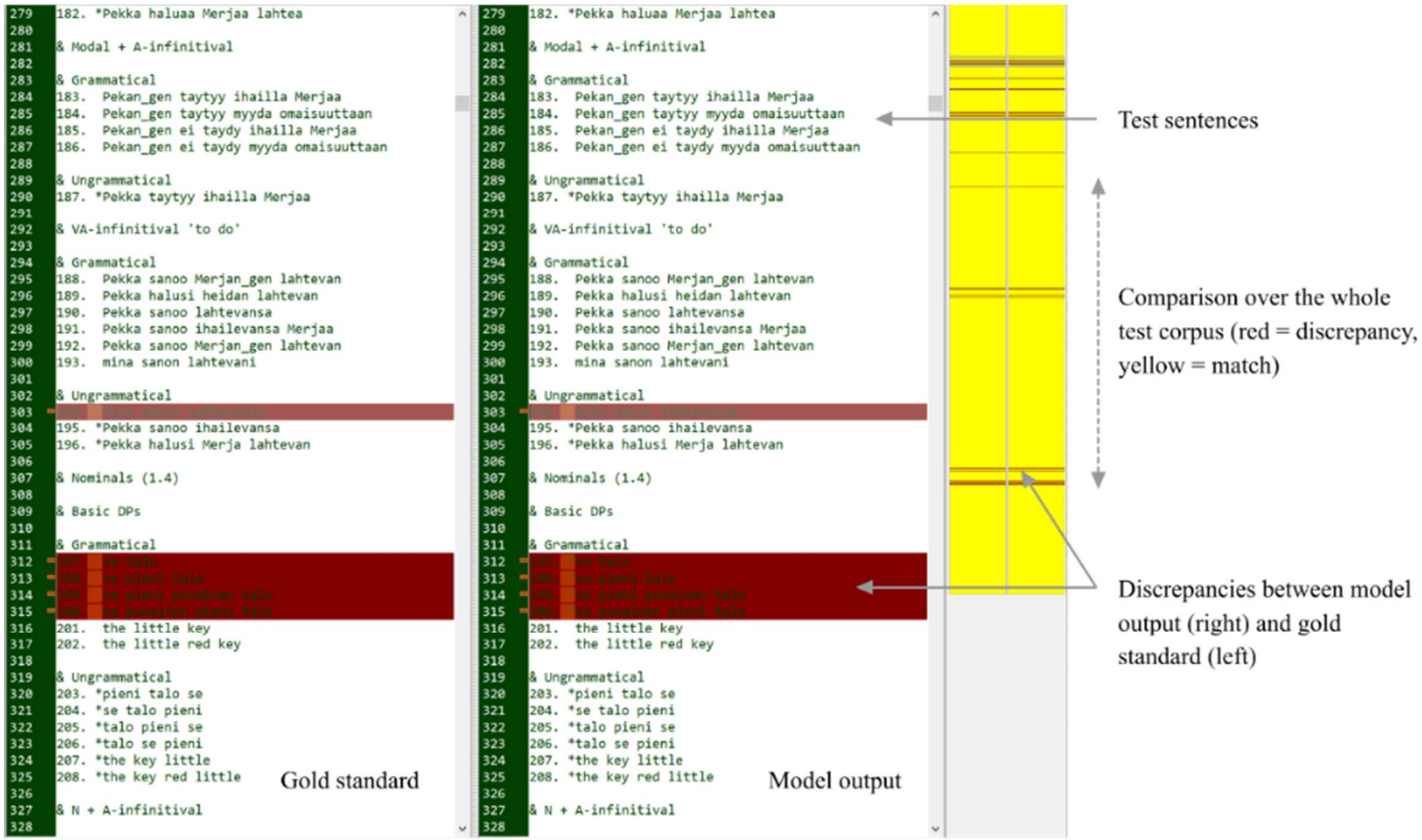


Figure 2: *A comparison between the gold standard (native speaker output, left) and calculated modal output (right). Discrepancies are marked by red lines by the automatic comparison tool. The rightmost yellow column shows the comparison over the whole test corpus*

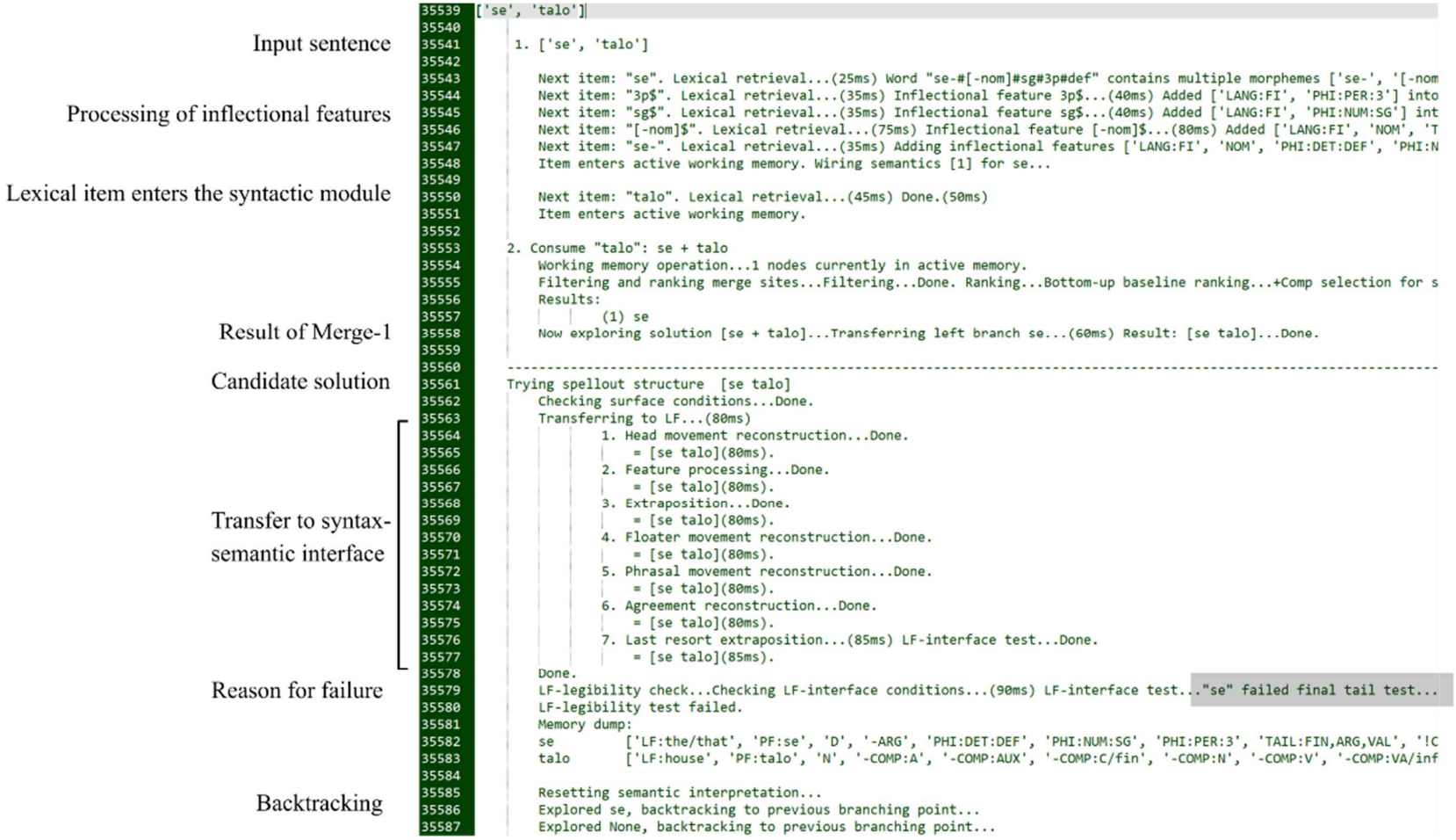


Figure 3: *Screenshot from the derivational log file showing the derivation of an isolated DP* se talo *'that.NOM house.NOM'*

The file is read in top-down order. The first element *se* 'the.NOM' is processed on Step 1 (lines 35541–35551), followed by the processing of *talo* 'house' (Step 2, lines 35553–35558). They are merged together to form [DP *se talo*] (line 35558), which is transferred (lines 35563–35578) to the syntax-semantic interface. The derivation fails at the syntax-semantics interface because the nominative case feature of *se* 'that.NOM' could not be checked: the required clausal context was missing (line 35579). The model tried to backtrack (lines 35586).

The hypothesis is revised until the model and data match. Once they do, the model is said to be *observationally adequate*. No natural language syllogisms or intuitive jumps occur in the justification, and no authority is allowed to use Augustine's divine illumination to consider that "nothing whatsoever is said by way of justifying this analysis" or that it does not make "any sense" or does not have "any application". In fact, any defect, problem or limitation is completely transparent. They are shown in Figure 2.

Suppose we have a hypothesis that is observationally adequate or nearly observationally adequate. This requirement alone is not sufficient, or even particularly interesting. All it says is that the algorithm captures something about the formal "shape" of the data. One can reach observational adequacy by storing the whole test corpus into the algorithm's memory. A trivial model of this kind does not contribute anything to any linguistic theory, in the same was as a table-lookup catalog of precomputed values or measurements used in engineering does not contribute anything to physics. Explanations which are relevant within the context of a linguistic theory are said to have *descriptive adequacy* (Chomsky 1965).

What counts as a descriptively adequate grammar depends on the linguistic framework. A linguist who has constructed an observationally adequate theory by using a connectionist model will judge the matter differently than another researcher who works with modern minimalism. Similarly, a linguist who begins from a semantic-based dependency grammar will have different concerns than one who adopts some variation of cognitive linguistics. The expert authority who judged the present model as grossly incorrect proposed that we should use a historical method to capture the same data. In general, then, descriptive adequacy is meaningful in relation to "grammars that are paired with some linguistic theory" (Chomsky 1964: 52). This is because it is "always possible to describe the linguistic intuition of the native speaker in a completely ad hoc way in any particular case if we drop the requirement that the grammar be constructed in accordance with some fixed model or if we allow the associated linguistic theory to be completely general and without content" (ibid.).

To assess descriptive adequacy, we examine the calculated output in the context of an existing linguistic theory. Let us consider the output in connection with an interrogative clause *ketä Pekka ihaile-e* 'who.PAR Pekka.NOM admire-PRS.3SG' (sentence #383 in the test corpus). The calculated output is shown in Figure 4. The input sentence is associated with a syntactic interpretation (line 9435) and illustrated further in Figure 5 generated by the algorithm.

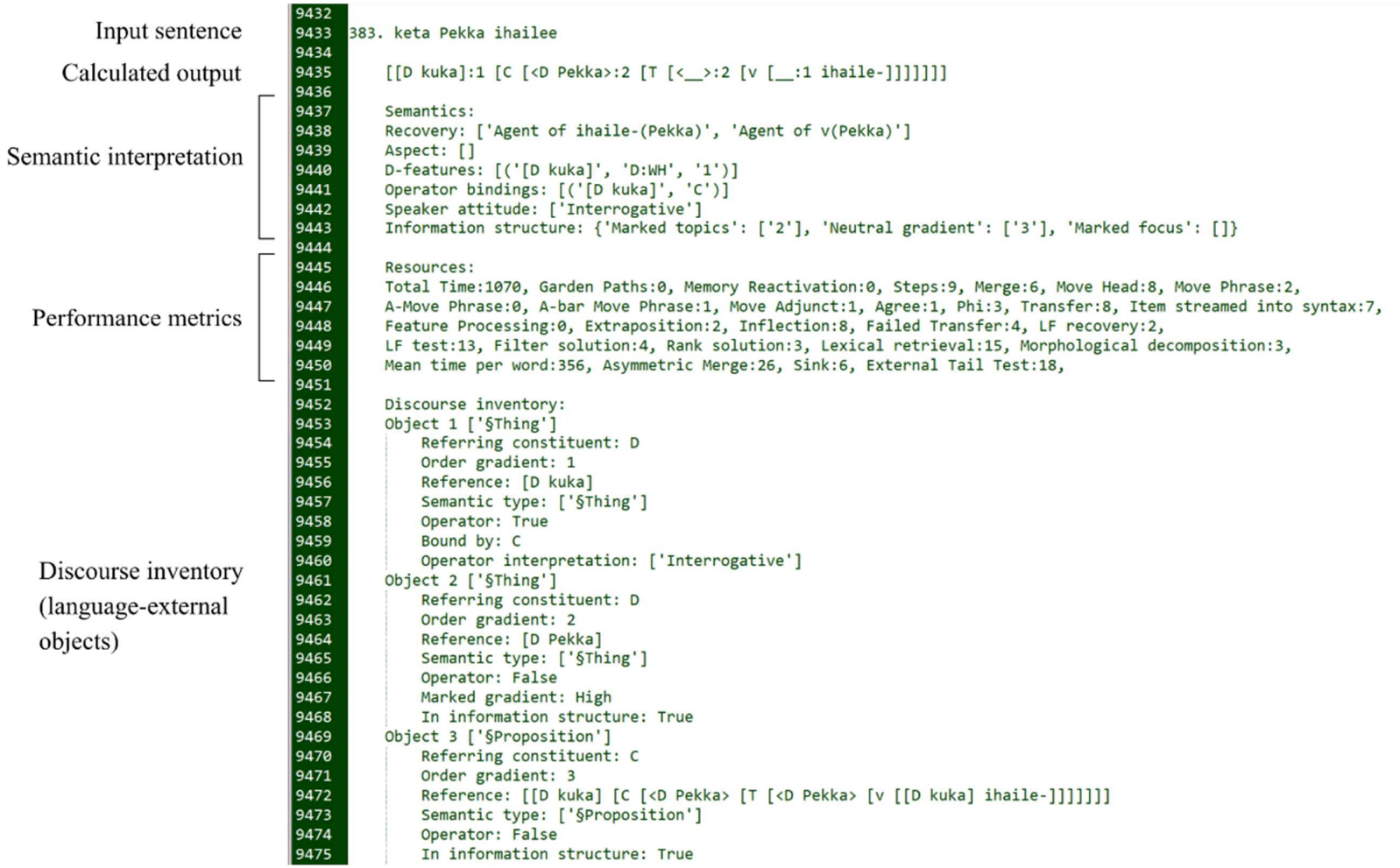


Figure 4: *Calculated output for* Ketä Pekka ihaile-e? *'who.PAR Pekka.NOM admire-PRS.3SG'*

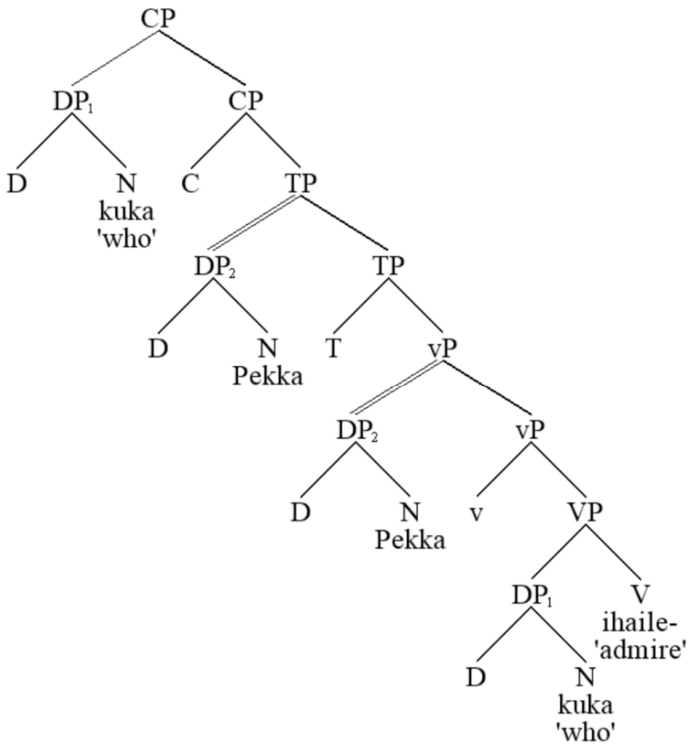


Figure 5: *The phrase structure analysis calculated by the model for the input sentence #383*

The researcher must determine whether the calculated syntactic interpretation is correct and/or plausible, given the background theory and the independent evidence that has motivated it. In this case it is: the interrogative pronoun was reconstructed correctly to the object position, while the grammatical subject/topic was reconstructed to the thematic SpecvP position. From the point of view of the theoretical framework used in this particular theory this is a plausible output.

The fact that descriptive adequacy must be relativized to a background theory or framework does not mean that no progress is possible beyond this point or that anybody can believe anything given any background theory. Let's consider, as an example, the fact that the algorithm analyzed the sentence as consisting of hierarchically organized parts. Suppose we change the algorithm so that it replaces hierarchical structures with flat representations. We will also write a function that interprets flat representations. What motivates any of these analyses if observational adequacy can be reached by either model?

The structure-dependence principle is based on linguistic evidence indicating that the object and the verb constitute one unit that does not include the subject. We could therefore claim that it is supported by empirical evidence. The problem, though, is that this justification is not visible in the calculations. Structure dependence can be justified for example by relying on reflexive binding, yet reflexive binding was not part of the algorithm and there were no reflexives in the test set. Therefore, the force one is willing to grant to this argument is to a degree subjective and, consequently, a significant portion of professional linguists remain skeptical towards these claims. One is free to assert by divine illumination or by some other superior cognitive ability that the same properties could be explained by a number of competing but simpler models, such as by purely historical analysis—or indeed that they could not.

Yet, we can now evade the whole issue. Suppose a researcher develops a model of Finnish word order that generates flat structures instead of hierarchical ones and provides a mechanism associating each flat structure with a semantic interpretation. If it reached observational adequacy and generated correct semantic interpretations, we would have two competing models that cannot be distinguished from each other in terms of their predictive success, one that uses hierarchical representations and another that projects flat structures. This would not be a problem, though, because eventually both researchers must face the evidence relevant for deciding the issue. We proceed by generating an agreed-upon dataset deemed relevant to the issue, thus it should contain at least reflexive data, control constructions, subject–object extraction asymmetries, and other data pertinent to this issue, and examine how each model handles that set. If the matter cannot be settled by this method, on the other hand, then the two models must be judged notational variants of only one underlying model—not an unusual situation and in no way a barrier to progress in the advanced sciences.[3]

Let us consider semantic interpretation next. Figure 6 shows part of the calculated semantic output for a simple interrogative clause.

```
9433 383. keta Pekka ihailee
9434
9435     [[D kuka]:1 [C [<D Pekka>:2 [T [<__>:2 [v [__:1 ihaile-]]]]]]]
9436
9437     Semantics:
9438     Recovery: ['Agent of ihaile-(Pekka)', 'Agent of v(Pekka)']
9439     Aspect: []
9440     D-features: [('[D kuka]', 'D:WH', '1')]
9441     Operator bindings: [('[D kuka]', 'C')]
9442     Speaker attitude: ['Interrogative']
9443     Information structure: {'Marked topics': ['2'], 'Neutral gradient': ['3'], 'Marked focus': []}
```

Figure 6: *Part of the calculated semantic output for a basic interrogative sentence. There is an error in the calculation; see the main text for explanation*

Line 9438 shows that the model interprets Pekka as the agent of the whole event ('who admires') but wrongly interprets Pekka as an argument of the verbal stem 'admire'. The latter should have been 'who'. The sentence does not mean 'Pekka admires himself'. Examination of derivations of other sentences in the test corpus reveals that this problem has to do with the fact that the patient is an interrogative operator: regular direct objects are interpreted correctly. But in this case the model output does *not* match with the semantic intuition of a native speaker, and some correction is needed. The rest of the interpretation appears to be correct.

The semantic interpretation shown in Figure 6 pools various aspects of semantic interpretation calculated during processing. The researcher can populate this structure with anything deemed relevant for a particular research agenda. Ideally, we would like to base the calculations on a more principled semantic system. Some initial steps towards such an explanation are taken by using a data structure called discourse inventory, shown in Figure 7.

[3] A well-known example of this is Freeman Dyson's proof of the equivalence of Feynman's and Tomonaga-Schwinger's approaches to quantum electrodynamics (Dyson 1949).

```
Discourse inventory:
Object 1 ['§Thing']
    Referring constituent: D
    Order gradient: 1
    Reference: [D kuka]
    Semantic type: ['§Thing']
    Operator: True
    Bound by: C
    Operator interpretation: ['Interrogative']
Object 2 ['§Thing']
    Referring constituent: D
    Order gradient: 2
    Reference: [D Pekka]
    Semantic type: ['§Thing']
    Operator: False
    Marked gradient: High
    In information structure: True
Object 3 ['§Proposition']
    Referring constituent: C
    Order gradient: 3
    Reference: [[D kuka] [C [<D Pekka> [T [<D Pekka> [v [[D kuka] ihaile-]]]]]]]
    Semantic type: ['§Proposition']
    Operator: False
    In information structure: True
```

Figure 7: *A screenshot of the output containing contents of the discourse inventory created during the processing of a simple interrogative clause*

The discourse inventory is populated with language-external semantic objects during the derivation. The term "semantic object" refers to (mental representations of) language external objects, such as persons, propositions or events that are denoted by the linguistic expressions that occurred in the input clause.

In short, then, descriptive adequacy is evaluated by comparing the model output against semantic and syntactic intuitions of native speakers and/or against a theoretical matrix as defined by a larger linguistic background theory.

Explanatory adequacy refers to a further requirement which demands that the model agrees with external evidence concerning language acquisition. Thus, a grammar that satisfies the condition of explanatory adequacy provides "a principled basis, independent of any particular language, for the selection of the descriptively adequate grammar of each language" (Chomsky 1964: 29), where "selection" refers, or can refer to, learning. A model of this type describes a language acquisition device that maps sensory data available to a language learner into a descriptively adequate grammar, relying on "innate predisposition of the child to develop a certain kind of theory to deal with the evidence presented to him" (Chomsky 1965: 26).

Any language comprehension algorithm will, as a matter of necessity and independent of whether it models language learning or not, contain fixed and variable parts. How the division is implemented is an empirical question. The algorithm used here assumes that while the computational principles remain universal, lexicons may differ. A model of this kind reaches explanatory adequacy if and only if it captures observations from several (or, in an ideal sense, all) languages. Explanatory adequacy can therefore be assessed by using several languages in the test corpus. In other words, such a theory must "develop an account of linguistic universals that […] will not be falsified by the actual diversity of languages" (Chomsky 1965: 28).

Psycholinguistic adequacy refers to the condition that the model should not contradict anything known independently concerning real-time language processing. To move towards such a goal, we can give the program an ability to monitor its own performance.

Thus, the results file contains a segment providing performance metrics, reproduced in Figure 8.

```
9445    Resources:
9446    Total Time:1070, Garden Paths:0, Memory Reactivation:0, Steps:9, Merge:6, Move Head:8, Move Phrase:2,
9447    A-Move Phrase:0, A-bar Move Phrase:1, Move Adjunct:1, Agree:1, Phi:3, Transfer:8, Item streamed into syntax:7,
9448    Feature Processing:0, Extraposition:2, Inflection:8, Failed Transfer:4, LF recovery:2,
9449    LF test:13, Filter solution:4, Rank solution:3, Lexical retrieval:15, Morphological decomposition:3,
9450    Mean time per word:356, Asymmetric Merge:26, Sink:6, External Tail Test:18,
```

Figure 8: *Performance metrics provided by the algorithm when processing a simple intransitive clause*

Total time refers to the predicted processing time in milliseconds, thus the model predicts that it takes an average hearer approximately one second to process the sentence *ketä Pekka ihaile-e?* 'who.PAR Pekka.NOM admire-PRS.3SG'. These numbers are created by associating the linguistically relevant computational operations postulated by the algorithm with a predicted processing cost in milliseconds and then summing them over the course of the whole derivation. Predicted costs should ultimately be determined on the basis of well-established physiological properties of human neuronal information processing and then tested in laboratory experiments. They can also be used to assess the computational cost and psycholinguistic reality of any proposed linguistic model.

Performance metrics are also provided in a comma-delimited form that can be processed by external programs such as Excel, SPSS, or by Python data processing scripts as I did below. The file lists each input sentence associated with the performance metrics shown above, all written to the same line. We can group the input sentences from this file on some basis, say by using the classification present in the test corpus in Table 1. These groups can then function as independent variables; for dependent variables, we take whatever metrics are of particular interest. Figure 9 shows the results when I examined the mean predicted processing times per word as a function of the main linguistic category in the test corpus.

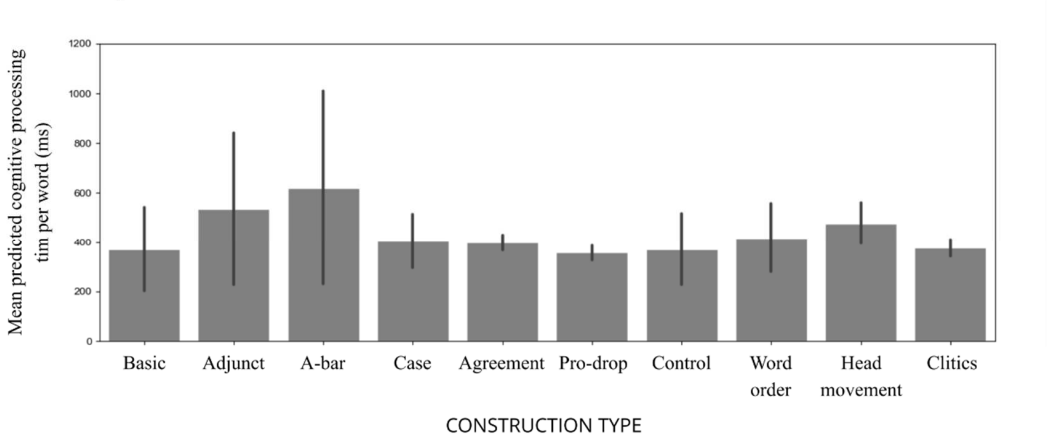


Figure 9: *Mean predicted processing times (in milliseconds) as a function of construction type. Errors bars represent standard deviation*

The first category "Basic" (Group 1 in the test corpus, Table 1) can be taken as an overall estimation of how the model performs with standard clause types. A word is predicted to take an average of 372 ms to process. The model has difficulties with sentences that involve

adjunction (mean 535 ms) and operator movement (mean 619 ms). This is due to garden pathing, as shown by Figure 10.

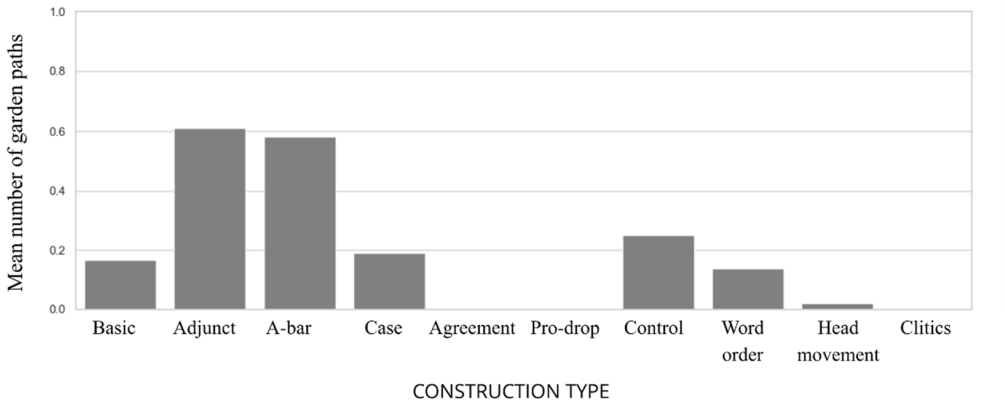


Figure 10: *Mean number of garden paths as a function of major category in the test corpus*

Whether performance properties of this kind are included into the study is for the researcher to decide.

4 Finnish agreement and computation

I will conclude this note by applying the methodology to the study of Finnish agreement to illustrate how it works in connection with a concrete empirical problem.

Finnish finite elements agree in number and person with nominative grammatical subjects (2a–e). The category of Finnish finite elements contains finite verbs (2a, e), auxiliaries (2b), negation (2c), and special modal verbs (2d).[4]

(2) a. *Jari* *ihaile-e* *naapure-i-ta-an.*
Jari.NOM admire-PRS.3SG neighbour-PL-PAR-PX/3SG
'Jari admires his neighbors.'

b. *Jari* *o-n* *ihail-lut* *naapure-i-ta-an.*
Jari.NOM be-PRS.3SG admire-SG.PST.PRTCPL neighbour-PL-PAR-PX/3SG
'Jari has admired his neighbors.'

c. *Jari* *e-i* *ihail-lut* *naapure-i-ta-an.*
Jari.NOM not-3SG admire-SG.PST.PRTCPL neighbour-PL-PAR-PX/3SG
'Jari did not admire his neighbours.'

d. *Naapure-i-den* *täytyy* *pitä-ä* *ovi* *lukossa.*
neighbour-PL-GEN must.0 keep-A/INF door in.locked
'The neighbours must keep the door locked.'

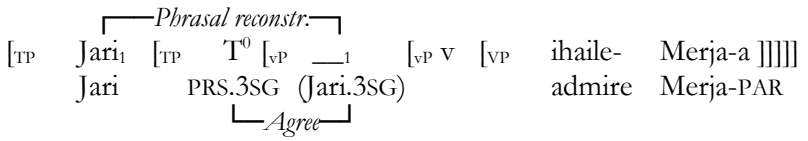
[4] Abbreviations: 0 = no agreement or default third person agreement; 1, 2, 3 = person features; A/INF = A-infinitival; GEN = genitive case; NOM = nominative case; PAST = past tense; PAR = partitive case; PL = plural; PRS = present tense; PRTCPL = participle; PST = past tense; PX = infinitival agreement marker; SG = singular.

e. *Mei-tä väsyttä-ä.*
we-PAR feel.tired-PRS.3SG
'We feel tired.'

The algorithm gets syntactic agreement configurations from the parser that it must check against possible phi-feature mismatches between the subject and finite element. This checking relation must involve finite elements and nominative thematic subjects (2a–c), and it must ignore non-nominative arguments, independent of their positions (2d–e). Checking cannot be performed presyntactically, it must access syntactic notions such as 'subject' and 'finite element'. In addition, it is not in my view possible to perform these operations with a partial phrase structure whose final structural properties remain unknown. Another option is to perform the required checking operations in some postsyntactic semantic component, but the empirical evidence (e.g., (2a–e)) indicates this to be unlikely. These considerations suggest that agreement checks are performed between the parsing stage and semantic interpretation, therefore during transfer. Accordingly, let us assume that transfer contains an agreement reconstruction operation which verifies that there are no agreement mismatches in the phrase structure produced by the parser and prior reconstruction operations.

To sharpen this idea into a formal model, we consider first a tentative agreement hypothesis that is close to the standard agreement operation assumed in virtually all current minimalist literature on Finnish (e.g., Koskinen 1998, Holmberg & Nikanne 2002, Manninen 2003, Brattico 2019b) and on UG (Chomsky 2000, 2001, 2008). According to this analysis, finite elements agree in number and person with a local DP at the thematic SpecvP position (3). The DP is reconstructed from the preverbal topic position SpecTP during transfer by a reverse-engineered phrasal movement called phrasal reconstruction.

(3)



We will need some way to alert agreement reconstruction that T^0 requires an operation of this type. The minimalist theory available at present uses special-purpose vehicles called "probe features" for this purpose (Chomsky 2000, 2001), so let's borrow this idea and assume that there is a feature, call it [VAL] from "valuation required", which triggers the operation. We express the hypothesis in Python and run it against the test set.[5] The results are shown in Figure 11.

[5] The assumptions stated in the main text are not sufficient to specify a full Python implementation. We need a mechanism for finding grammatical heads with [VAL] from the structure, since no formal algorithm can "eyeball" them. In addition, we must specify what the local configuration between T and DP is, what will happen when Agree occurs, what happens if there are several arguments at the edge of vP, how DP adverbials are separated from arguments, and so on. Some of these issues will be examined below.

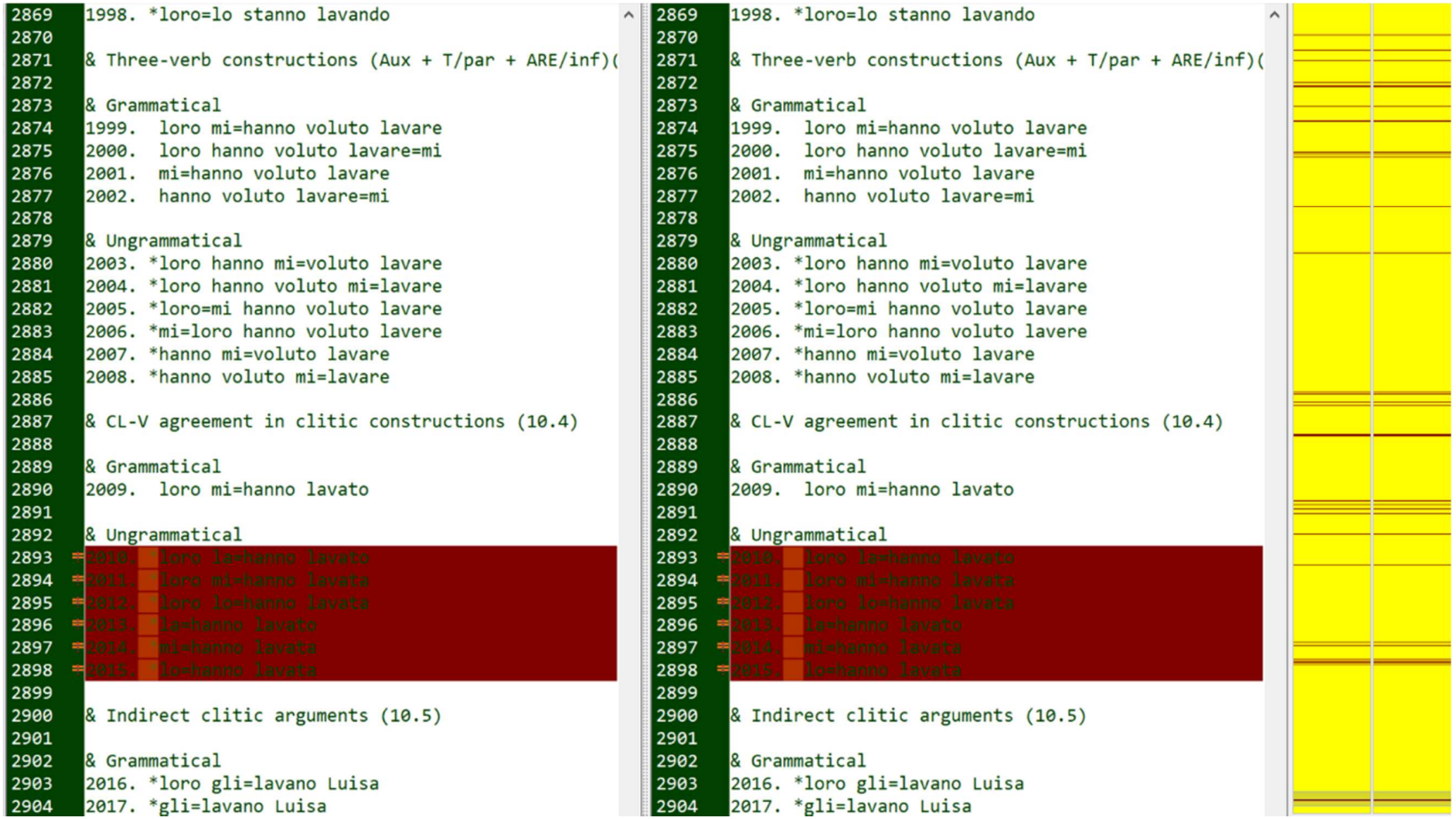


Figure 11: *Calculated results from the first trial simulation*

Several errors emerged (compare Figure 11 with Figure 2). One clear cluster is Italian clitic agreement (lines 2893–2898 in Figure 11). Properties of Italian clitic constructions indicate that there must exist an agreement dependency between a head and its specifier (i.e. clitic–participle agreement), which (3) ruled out. I return to this below; for the time being we consider Finnish. By examining the discrepancies in the Finnish datapoints we find, first, that the model is unable to handle *pro*-drop constructions. The subject pronoun can remain null in Finnish if it is not in the third person (4a–b) (Vainikka 1989, Vainikka & Levy 1999, Holmberg 2005, Holmberg & Sheehan 2010).

(4) a. *Siivo-sin koko päivän.*
clearn-PST.1SG all day
'I cleaned all day.'
b. **Siivo-si koko päivän.*
clean-PST.3SG all day
'He cleaned all day.'

First, though, we must note that (3) becomes vacuous under these circumstances. There is no overt argument against which the features of the verb could be checked. In addition, it is easy to image structural analyses under which the verb would wrongly agree with the DP adverbial 'whole day'. Let us fix this by reinterpreting the theory so that it only requires 'no mismatches', hence positive matching is not required. Sentences such as (4a) are now correctly judged grammatical. The problem, though, is that (4b) passes as well.

Here we have to consider an additional factor. The presence of agreement features seems to license the *pro*-drop phenomenon. If we licensed *pro*-drop everywhere without taking agreement into account, then all hypothetical English *pro*-drop sentences such as **admires Mary* would be wrongly calculated to be grammatical. To capture the contrast between (4a–b) while rejecting English *pro*-drop sentences across the board, it is usually assumed in the linguistic literature that the Finnish third person agreement features as well as English agreement are in some sense too 'weak' to license *pro*-drop. Perhaps we can think of strong agreement clusters as replacing overt pronoun subjects, in some sense. We could easily draw a distinction between weak and strong agreement markers in the lexicon. What complicates the issue is that sentences like (4b) *are* accepted in Finnish if they occur in a context where the missing third person subject can be linked with an acceptable antecedent (5).

(5) *Pekka väitti että* [*siivo-si koko päivän*].
Pekka claimed that clean-PST.3SG all day.
'Pekka claimed that he (= Pekka) cleaned all day.'

Agreement features license *pro*-drop in Finnish if and only if they are in the first or second person (4a–b) *or* they are in the third person and there is an antecedent (5). The pattern is summarized in Table 2.

	Person	
	1st, 2nd	3rd
Antecedent		
No	Yes	No
Yes	Yes	Yes

Table 2: *Licensing of pro-drop in Finnish as a function of agreement and antecedent*

At this point our analysis needs a mechanism for finding the required antecedents, but I will omit this issue here; see Brattico (2021), which follows the Holmberg–Sheehan hypothesis discussed further below. Test simulation shows that the proposed mechanism is still insufficient, however, because it accepts verb-initial sentences such as (6).

(6) **Siivo-si Pekka koko päivän.*
clean.PST.3SG Pekka.3SG all day
└─Agree─┘

T^0 agrees with the postverbal subject, but the clause is not grammatical. It cannot be just the weakness of the third person agreement that matters. The most likely reason for the ungrammaticality of (6) is that verb-initial clauses are generally ungrammatical in Finnish. In the generative theorizing this generalization is captured by stipulating that finite elements contain a topic-based "EPP-feature" that must be checked by a subject/topic phrase (Vainikka 1989, Vilkuna 1995, Holmberg & Nikanne 2002, Brattico 2019b, Huhmarniemi 2019). The sentence is grammatical if some phrase, typically but not necessarily the grammatical subject, is moved to the preverbal topic position. What we must assume, it seems, is that the "strong" first and second agreement features suffice to check this EPP condition, while third person features do not. There is a connection between agreement and the presence of a phrase at the local specifier position, and indeed now we recall that it was the absence of the same spec–head configuration that caused the original problem with the Italian clitic data. Not even this would be sufficient, however, since third person agreement features *do* suffice to the check the EPP if (and only if) there is the antecedent (see (5) and Table 2).

To me, the relativization of Finnish EPP to the presence of an antecedent has always represented something of an enigma. I will consider one possible avenue that I have used to obtain best results so far, although many problems remain. Following Holmberg (2005) and Holmberg and Sheehan (2010), we consider that the Finnish third person agreement features, in contrast to 1st and 2nd person, are not strong enough to check the D-feature of the finite element, which triggers antecedent search as a last resort, capturing the Finnish partial *pro*-drop signature. This mechanism was implemented into the algorithm reported in Brattico (2021). To connect the mechanism to the Finnish EPP and Italian clitic agreement data, we assume that agreement with the D-feature relies on the spec–head configuration, which then replaces the EPP requirement. Hence Finnish EPP = checking of a nominal D-feature, generating a definiteness effect instead of a topic effect (Brattico 2019b). Furthermore, we generalize agreement so that it checks elements both inside the sister and specifier of the triggering head. Test simulations showed that the Italian clitic data now comes out correctly. The Finnish EPP data follows if we assume that D-checking from the specifier position is mandatory: first and second person verbs can remain without overt subject, being strong enough to check the D-feature all by themselves, while third

person verbs are too weak to check the D-feature, requiring an overt phrase at SpecTP or antecedent support. This agrees with the claim made by Holmberg and Nikanne (2002) that the Finnish topic EPP can be checked by any phrase that is "broadly referential", thus we interpret this claim as requiring that there is a D-feature inside the phrase occupying the specifier position. Third person constructions are grammatical if the D-feature is valued by antecedent recovery. The code corresponding to the revised final model is shown in Figure 12.

```
1   def Agree_1(self, head):
2
3       goal, phi_features = self.Agree_1_from_sister(head)
4       for phi in phi_features:
5           self.value(head, goal, phi, 'sister')
6       if not head.is_unvalued():
7           return
8
9       goal, phi_features = self.Agree_1_from_edge(head)
10      for phi in phi_features:
11          if find_unvalued_target(head, phi):
12              self.value(head, goal, phi, 'edge')
```

Try to agree with elements inside sister constituent

If agreement leaves features unvalued, try to agree with specifier

Figure 12: *Python implementation of the final version*

Running the test corpus with this model shows that although the situation improves, the hypothesis fails to handle (7), which it judges wrongly as ungrammatical.

(7) *Merja-a* *ihaile-n.*
Merja-PAR admire-1SG
'I admire Merja.'

Because the thematic subject is covert, T^0 does not find anything from its sister and targets the partitive DP at the specifier position. Because both phi-features and the D-feature remains unvalued, spec–head agreement is triggered, which produces feature mismatch. The problem is that only nominative arguments are relevant for agreement in Finnish. In the first iterations of the model this was not an issue, because the model targeted the SpecvP position that contained either nothing or the reconstructed subject, hence we got the nominative/agreement correlation for free. But as soon as we include specifier agreement into the model, the problem of correlating agreement with nominative arguments resurfaced.

How to proceed from here is empirically unclear, but methodologically straightforward. We craft a representative test corpus that captures the whole Finnish agreement signature, possibly in conjunction with English agreement for comparison, and write a model that calculates it. The test corpus must contain at least finite clauses with and without agreement, with and without *pro*-drop, with and without antecedents, with and without preverbal phrases, and all these in many or perhaps in all possible word orders. More examples could be added, of course, if deemed relevant. This will ensure that, if nothing else, at least our hypothesis cannot be deemed "very far removed" from what counts as a valid hypothesis simply by subjective opinion.

5 Conclusions

One possible reason for the lack of progress in the less advanced sciences such as linguistics could be their stubborn use of an antiquated research method in which theories are justified by relying on some form of Augustine's divine illumination, a supreme cognitive capacity accessing the veridicality of an ambiguously formulated idea by intuition, common sense or thought experiment. Although this method can produce interesting data and theories, it is unable to produce agreement, creating an obstacle for progress. The notion of rigorous proof was considered a possible solution.

References

Brattico, Pauli. 2019a. *A computational implementation of a linear phase parser. Framework and technical documentation (Version 9.0)*. Pavia: IUSS University School for Advanced Studies.

Brattico, Pauli. 2019b. Subjects, topics and definiteness in Finnish. *Studia Linguistica* 73. 1–38. https://doi.org/10.1111/stul.12129

Brattico, Pauli. 2020. Finnish word order: Does comprehension matter? *Nordic Journal of Linguistics* 44. 38–70.

Brattico, Pauli. 2021. Null arguments and the inverse problem. *Glossa: A Journal of General Linguistics* 6. 1–29. https://doi.org/10.1017/S0332586520000098

Chomsky, Noam. 1957. *Syntactic structures*. The Hague: Mouton. https://doi.org/10.1515/9783112316009

Chomsky, Noam. 1964. *Current issues in linguistic theory*. Berlin: Mouton. https://doi.org/10.1515/9783110867565

Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge, MA.: MIT Press.

Chomsky, Noam. 2000. Minimalist inquiries: The framework. In Roger Martin, David Michaels & Juan Uriagereka (eds.), *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*, 89–156. Cambridge, MA.: MIT Press.

Chomsky, Noam. 2001. Derivation by phase. In Michael Kenstowicz (ed.), *Ken Hale: A life in language*, 1–37. Cambridge, MA.: MIT Press.

Chomsky, Noam. 2008. On phases. In Carlos Otero, Robert Freidin & Maria-Luisa Zubizarreta (eds.), *Foundational issues in linguistic theory: Essays in honor of Jean-Roger Vergnaud*, 133–66. Cambridge, MA.: MIT Press. https://doi.org/10.7551/mitpress/9780262062787.001.0001

Dyson, Freeman J. 1949. The radiation theories of Tomonaga, Schwinger, and Feynman. *Physical Review* 75. 1–32. https://doi.org/10.1103/PhysRev.75.486

Holmberg, Anders. 2005. Is there a little pro? Evidence from Finnish. *Linguistic Inquiry* 36. 533–64. https://doi.org/10.1162/002438905774464322

Holmberg, Anders & Nikanne, Urpo. 2002. Expletives, subjects and topics in Finnish. In Peter Svenonius (ed.), *Subjects, expletives, and the EPP*, 71–106. Oxford: Oxford University Press.

Holmberg, Anders & Sheehan, Michelle. 2010. Control into finite clauses in partial null-subject languages. In Theresa Biberauer, Anders Holmberg, Ian Roberts & Michelle Sheehan (eds.), *Parametric variation: Null subjects in minimalist theory*, 125–52. Cambridge: Cambridge University Press. https://doi.org/10.1017/cbo9780511770784.004

Huhmarniemi, Saara. 2019. The movement to SpecFinP in Finnish. *Acta Linguistica Academica* 66. 85–113. https://doi.org/10.1556/2062.2019.66.1.4

Koskinen, Päivi. 1998. *Features and categories: Non-finite constructions in Finnish*. PhD dissertation, University of Toronto.
Manninen, Satu. 2003. *Small phrase layers: A study of Finnish manner adverbials*. Amsterdam/Philadelphia: John Benjamins. https://doi.org/10.1075/la.65
Phillips, Colin. 1996. *Order and structure*. PhD dissertation, MIT.
Vainikka, Anne. 1989. *Deriving syntactic representations in Finnish*. PhD dissertation, University of Massachusetts Amherst.
Vainikka, Anne & Levy, Yonata. 1999. Empty subjects in Finnish and Hebrew. *Natural Language & Linguistic Theory* 17. 613–71. https://doi.org/10.1023/a:1006225032592
Vilkuna, Maria. 1995. Discourse configurationality in Finnish. In É. Kiss Katalin (ed.), *Discourse Configurational Languages*, 244–68. Oxford: Oxford University Press.

Pauli Brattico
Independent Researcher
pauli.j.brattico@gmail.com