GREATER THAN ZERO?
A STUDY OF REFERENTIALLY SPECIFIC
AND OPEN NECESSITY CONSTRUCTIONS
IN FINNISH EVERYDAY CONVERSATION

Mikael Varjo
University of Turku, FI
mavarj@utu.fi

Abstract. This article examines unipersonal necessity constructions with *pitää* 'must, have to' in Finnish everyday conversation. The necessity constructions are unipersonal, which means that their predicate verb does not show agreement with the person of the subject. Omitting the pronominal subject altogether is common in necessity constructions, typically resulting in a so-called zero-person construction, which is a type of referentially open personal construction in Finnish. The present article sets out to explore the semantico-grammatical and discourse features of necessity constructions without an overt grammatical subject, comparing them with those that have explicit subjects (both speech-act pronouns and third-person subjects) by employing a generalized linear mixed model (GLMM). Following an overview of the statistical results, the article provides a qualitative analysis of the statistically significant differences. The findings indicate that the zero-person constructions of necessity have distinct functions in organizing social interaction, such as (joint) planning, expressing wishes, and generally constructing involvement between participants, whereas the necessity constructions that have a speech-act pronoun as the subject are more oriented toward emphasizing the necessity or obligation of taking an action rather than directly implementing it.

Keywords: corpus linguistics, Finnish language, interactional linguistics, necessity construction, open reference, syntax

1.   Introduction

This paper deals with the Finnish modal constructions of necessity (henceforth necessity construction or NEC when referring to this study's data set) with *pitää* 'must, have to'[1] and their use in everyday

---

1   *pitää* may express dynamic, deontic, and epistemic necessity. For a more extensive overview of the different interpretations, see Kangasniemi (1992).

conversation. Its focus lies on the differences between referentially specific and open necessity constructions. The referentially specific constructions consist of both speech-act pronouns as subjects and third-person pronouns or NPs as subjects, whereas the open necessity constructions in this study refer to the Finnish zero-person construction, an open personal construction in which the predicate verb is in the third-person singular form and there is no grammatical subject. The referent of the unexpressed subject is left open, and the construction can be used either to refer to different speech act participants or to make generalizations (e.g. Hakulinen & Karttunen 1973, Laitinen 2006). The aim of this study is to analyze the interactional tasks and the contexts of use of the referentially open necessity constructions compared with the referentially specific constructions. A generalized linear mixed model (GLMM) is employed to examine the semantico-grammatical and discursive differences between the referentially specific and open necessity constructions, and these differences are then further analyzed on the interactional level.

Necessity constructions typically express obligation or recommendation. In Finnish, these constructions are unipersonal, meaning that the modal verb is in the third-person singular, and they exemplify different patterns and degrees of grammaticalization among the modal verbs[2] (see also Helasvuo & Laitinen 2006 on the person marking in Finnish). In the unipersonal constructions, the grammatical person can only be expressed by means of personal pronouns or other subject NPs. This means that in necessity constructions, speakers have fewer linguistic devices to differentiate the personal forms and the zero person.

Examples (1a)–(1f) demonstrate the Finnish person marking and zero person. In (1a), the verb *mennä* is in the singular first-person form, and therefore, the pronominal subject could be omitted, whereas (1b) then illustrates the use of zero person in a similar context. Here, the 'if–then' frame is required for the zero person to be used with an

---

2   In the necessity constructions in Finnish, the subject is typically in the genitive (e.g. Helasvuo & Vilkuna 2008: 220), whereas other modal verbs have full person inflection, and the grammatical subject in these constructions is in the nominative (i.e., the prototypical subject in Finnish). Although the verbs that are used in the unipersonal constructions sometimes only have third-person forms (e.g. *täytyä*), many of them also have other than modal meanings compatible with the full person inflection (e.g. *pitää*, *tarvita*, *kannattaa*) in contemporary Finnish (see Kehayov & Torn-Leesik 2009: 366–367).

activity verb (Laitinen 2006: 212), and in contexts like this, the zero person is typically generic (e.g. Jokela & Plado 2015: 86). Example (1c) demonstrates how even the modal verbs that do not express necessity in Finnish agree with the subject of the clause in person and number (see Kehayov & Torn-Leesik 2009: 366). Again, the pronominal subject could be omitted. Example (1d) then illustrates zero person in a modal construction that does not express necessity, this time with a preverbal theme (Laitinen 2006: 213), as it states a modal generalization that there are no external circumstances hindering the recipient (or the speaker) from going to the grocery store (cf. Jokela & Plado 2015: 86). Lastly, examples (1e) and (1f) show how the unipersonal modal constructions contrast with the ones in (1c) and (1d): As there is no verbal subject marker in the unipersonal construction in (1e), omitting the pronominal subject would result in a zero-person construction.[3]

(1)   a.   *(Minä)*   ***menen***   *tänään*   *ruokakauppaan.*
            I   go.3SG   today   grocery.store.ILL
            'I (will) go to the grocery store today.'

      b.   *Jos*   ***menee***   *tänään*   *ruokakauppaan,*   *huomenna*
            if   go.3SG   today   grocery.store.ILL   tomorrow
            jääkaapissa   *on*   *ruokaa.*
            fridge.INE   is.3SG   food.PART
            'If Ø goes to the grocery store today, there is food in the fridge tomorrow.'

      c.   *(Minä)*   ***voin***   *mennä*   *tänään*   *ruokakauppaan.*
            I   can.1SG   go.INF   today   grocery.store.ILL
            'I can go to the grocery store today.'

      d.   *Tänään*   ***voi***   *mennä*   *ruokakauppaan.*
            today   can.3SG   go.INF   grocery.store.ILL
            'Ø can go to the grocery store today.'

---

3   Sometimes even a slight change in the word order can differentiate zero person from a personal form (e.g. an anaphoric zero, see Laitinen 2006: 213). In (1f), a prototypical zero-person construction would have the adverb *tänään* 'today' as a preverbal theme. Nevertheless, the open interpretation is still the most conventionalized one here as well, albeit somewhat dependent on the context.

e. ***Minun   pitää***   *mennä   tänään   ruokakauppaan.*
   I.GEN   must.3SG   go.INF   today   grocery.store.ILL
   'I have to / must go to the grocery store today.'

f. ***Pitää***   *mennä   tänään   ruokakauppaan.*
   must.3SG   go.INF   today   grocery.store.ILL
   'Ø has to / must go to the grocery store today.'

Adopting a usage-based approach (e.g. Hopper 1987, Bybee & Thompson 1997) and combining both quantitative and qualitative methods, it becomes possible to explain the variation in the necessity constructions in a systematic manner and to reflect upon the discourse-pragmatic functions they can be assigned to in talk-in-interaction. Employing a statistical model is useful to find recurring patterns in a language or language variant but also when trying to provide more information about the variables that explain the variation to make comparisons between languages and language types possible. Hence, the current study also provides a statistically oriented point of reference to recent findings regarding open person constructions in other Finnic languages (e.g. Uusitupa 2021).

2.   Finnish person reference and the relevance of the Finnish zero person

The Finnish person system consists of three persons in both the singular and plural, and the predicate verbs tend to agree with nominative subjects in person and in number. From a typological perspective, Finnish can be regarded as a partial "pro-drop" language as the first and second person pronominal subjects are typically optional (e.g. Heinonen 1995, Dryer 2013). There are, however, differences between written and spoken interaction. Whereas double marking of the subject with both a pronominal and a verbal subject marker (i.e., inflectional suffix in the verb denoting the person and number) is the preferred alternative in conversational language, the written varieties (both standard and more colloquial) more often avoid pronominal subjects unless they serve a specific discourse function (e.g. Helasvuo & Laitinen 2006: 179, Helasvuo 2014, Helasvuo & Kyröläinen 2016: 266, Väänänen 2016).

The necessity constructions, however, represent an exception as their predicate verbs are unipersonal third-person singular forms, and

the personal pronouns in the subject role are in the genitive. If the clause is transitive, this applies to the third-person forms as well (both pronominal and other subject NPs). If intransitive, the subject is typically in the genitive if the referent of the subject is human or otherwise animate and is treated as an intentional being (see Helasvuo & Laitinen 2006: 189). These unipersonal constructions have also been referred to as impersonal modal patterns as they do not trigger agreement with the (agentive) nominal argument (Kehayov & Torn-Leesik 2009: 366), although the concept of unipersonality can be regarded as more strictly oriented to grammatical form than impersonality (see Helasvuo & Vilkuna 2008: 219–220). As Kehayov and Torn-Leesik (2009: 366–370) point out, the patterns regarding the subject–verb agreement vary across the modal verbs in the Balto-Finnic languages. Similar patterns exist in other European languages as well: Zinken and Ogiermann (2011), for example, have explored the Polish impersonal modal declarative of *trzeba* 'one needs to', which cannot receive a grammatical subject in any circumstances. The difference between *trzeba* and other modal requests is that whereas the latter focus on the recipients' ability and willingness, the former emphasize the participants' understanding of the requirements of the situation and are used to build shared responsibility between the interlocutors for initiating the action (Zinken & Ogiermann 2011: 284).

The so-called zero-person construction in Finnish consists of a third-person singular verb form without the lexical subject. The term zero subject has been used when the perspective is strictly limited to zeroes in the subject position (e.g. Laitinen 2006: 213). Zero person typically creates a deictically open reference and thus bears some resemblance to, for instance, the impersonal man construction in the Germanic languages (e.g. Nielsen et al. 2009, Ragnarsdottir & Strömqvist 2005). Zero-person constructions also resemble some third-person singular impersonals in Romance and Slavic languages (e.g. Siewierska 2008: 18–21). These reflexive impersonals, as Siewierska (2008: 18–21) calls them, share many common features with zero-person constructions: Their referents are necessarily human, they can be formed from both transitive and intransitive verbs, and, referentially, they tend to denote people in general, although they often also include the speaker and/or the addressee (Hakulinen & Karttunen 1973: 157−158, Laitinen 2006: 212, 218).

In Finnish, necessity constructions do not typically have a subject (in other words, they have a null or zero subject, e.g. Heinonen 1995, Laitinen 2006): For instance the verb *pitää* 'must, have to' appeared with a zero subject in 65% of the cases in the Finnish Parole Corpus[4] (Hakulinen et al. 2004: §1354). The tendency for the zero person to favor modal verbs has been widely recognized (e.g. Laitinen 1995, 2006; see Jokela 2012: 37–40 for a comprehensive overview). Statistically analyzing the zero-person constructions has also increased knowledge of the functions of the zero-person constructions that contain a modal verb in organizing social interaction: They are typically used to comment on something that has been discussed but also as proposals and directives. Because of their open reference, they are a common resource for conveying generalizations regardless of whether the speaker has firsthand experience. As such, they frequently function as expressions of personal stance (Varjo 2019: 74–77).

In Estonian, it is also common for zero subjects and modal verbs to co-occur, although the Estonian zero person in general has more restricted contexts of use than its Finnish counterpart. Comparing translations of written Finnish and Estonian has revealed that, besides the Estonian zero person, the equivalents for the Finnish zero person in different grammatical contexts include *da*-infinitive, the generic second-person singular, and the Estonian impersonal (Jokela & Plado 2015: 85–86).

Referentially open personal constructions, modal or not, commonly appear in contexts of planning (e.g. Pajusalu 2015: 47, Couper-Kuhlen & Etelämäki 2015: 16–19), are used as directives or suggestions (e.g. Jokela & Plado 2015: 88, Zinken & Ogiermann 2011), or express personal wishes, also when initiating others to join (e.g. Varjo 2019: 74–75). As such, they provide a speaker with resources for organizing and shaping the ongoing interaction while avoiding reference to person: The reference and the planned action are construed together in the context by the participants, and, in this way, all participants have equal access to the situation. The necessity constructions, on the other hand,

---

4   The Parole Corpus consists mainly of written texts and does not in that regard compare to the Arkisyn corpus used in this study. However, the zero-subject occurrences make up the majority of the necessity constructions (approximately 56%) in the data set of this study as well. The exact numbers of occurrences are mentioned in Section 3.

are commonly used as recommendations or obligations. The focus of this study is to find out when and why those recommendations and obligations are referentially open.

3.   Corpus and methods

3.1. Data and coding

The study data come from the Arkisyn database of conversational Finnish, a morphosyntactically annotated corpus that consists of approximately 30 hours of everyday conversations, both face-to-face and on the telephone. All of the conversations represent informal spoken Finnish. The conversations originate from the Conversation Analysis Archive at the University of Helsinki and the Finnish Language Recording Archive at the University of Turku. In its current form, Arkisyn contains altogether 279,023 words, making it a unique corpus of contemporary spoken Finnish in both its size and annotation (see Arkisyn 2018).

All the occurrences of NECs with the verb *pitää* were first extracted from the corpus, and the annotation of the corpus was then utilized to exclude clauses that did not belong to the scope of this study, such as existential constructions (e.g. *Täällä pitäisi olla enemmän ihmisiä.* 'There should be more people here') and certain other more marginal clause types (e.g. Hakulinen et al. 2004: §891). In total, 637 NEC occurrences with *pitää* as their main (i.e., finite yet unipersonal) verb met the requirements.[5] Two other verbs of necessity, *täytyä* and *tarvita*, were considered for the study but were eventually left out because they did not provide enough data points for the statistical analysis.[6]

3.2. Statistical analysis

In the statistical analysis, a GLMM was employed in R (R Core Team 2020) with functions from package lme4 (Bates et al. 2015). GLMM was used to perform two comparisons, both of which had the

---

5   Data is available at https://osf.io/uf8c5/.

6   In the Arkisyn database, the absolute frequencies of the verbs of necessity are as follows: *pitää,* N = 700; *täytyä,* N = 173; *tarvita,* N = 141.

realization of the subject NP of the clause as the dependent variable. As this study aims to reveal the differences between the open necessity constructions and the referentially specific constructions, the comparisons are made separately instead of a single multinomial model. This choice also makes the interpretation of the results easier (see e.g. Gries 2021: 344-353 about interpreting multinomial models and the complexity and pitfalls related). In the first comparison, the NECs with no overt subject NP (henceforth CovertSubj) were set against those with a speech act pronoun as their (genitive) subject (henceforth SpeechActSubj). The second comparison was between CovertSubj and NECs with a third-person NP as the (genitive) subject (henceforth 3PersSubj). The groups were adjusted to be of equal size in the comparison: The sample size of the largest group (CovertSubj, N = 354) was randomly adjusted so that it matched that of SpeechActSubj (N = 189) and 3PersSubj (N = 94).

GLMM was chosen due to its ability to provide estimates based on the interactions between the fixed (explanatory) effects, as well as to consider random effects, meaning that both by-item and by-subject variation are accounted for (e.g. Winter 2013: 5-6). Thus, it is possible to avoid the "language-as-fixed-effect fallacy" (see Clark 1973), according to which the failure to indicate items as a random factor can lead to serious error. The statistical significance of each fixed variable was then measured by means of a likelihood ratio test to compare the model that included an individual variable with a model that did not include that individual variable (see Gries 2021: 267–269). In this paper, the p-values below 0.05 were considered to indicate statistically significant effects.

The type of the subject NP was modelled as a function of a set of fixed effects based on earlier research on the Finnish zero person, such as adverbials in the initial field[7] and the presence and location of object NPs, conditional mood, presence of discourse particles, and past tense (e.g. Laitinen 1995, 2006, Hakulinen et al. 2004: §1347–§1356, Varjo 2019). *että*-clauses (Koivisto, Laury & Seppänen 2011) were included in the model to portray the sequential structure of conversation and to

---

7   I use the term *initial field* when referring to the syntactic position of the preverbal elements. This is based on the so-called field description of the Finnish word order (Vilkuna 1989, Hakulinen et al. 2004: §1369), where the term *alkukenttä* 'initial field' includes both *esikenttä* 'pre-field' and *teemapaikka* 'theme position'.

consider the resources the interlocutors have to organize the participant framework (Goodwin 1984, Goodwin & Goodwin 1990). All the fixed and random effects along with their possible values are introduced in Table 1.

Position of the clause inside a turn was the only sheer discursive variable, and it was used to monitor the positioning of the utterance within one speaker's turn: the clause may be simultaneously turn-initial and turn-final, either turn-initial or turn-final or neither (i.e. in the middle of the turn). Position in the turn is thus linked to the structure of conversation and narrativity as it indicates whether speaker changes occur nearby.

Although verb semantics per se cannot be considered in this study because the predicate verb is always the same, the infinitival complements of *pitää* are included in the analysis as a random effect to examine how the relationship between the dependent variable and the fixed effects varies across the infinitival complements (e.g. Winter & Grice 2021). In addition to the infinitival complements, speaker was included in the analysis as a random effect to control the idiosyncratic variation due to individual differences (Winter 2013: 2). In the case of dummy variables, the absence of the feature is used as the reference group. With object NP, the reference group is position other than the initial field, and with position of the clause inside the turn, it is those clauses that are simultaneously turn-initial and turn-final.

Table 1. Fixed and random effects used in the GLMM.

Fixed effects	Values	Example
Semantico-grammatical		
Adverbial in the initial field	yes/no	***Huomenna*** *pitää herätä aikaisin* tomorrow must.3SG wake.up early 'Ø has to wake up early tomorrow'
Adverbial conjunction	yes/no	***Jos*** *pitää kuitenkin vuorotella* if must.3SG nevertheless take.turns.INF 'If Ø nevertheless has to take turns'
että-clause	yes/no	***Että*** *pitää kaikki dokumentoida* COMP must.3SG everything document.INF '(So that) Ø has to document everything'

Fixed effects	Values	Example
Conditional mood[8]	yes/no	*Se* **pitäisi** *tehdä pian* DEM must.3SG do.INF soon *'Ø should do that soon'*
Interrogative form	yes/no	*Pitäisi**kö** se vaihtaa tuohon?* must.3SG.COND.Q DEM change.INF DEM.ILL *'Should Ø change it into that?'* **Miksi** *huomenna pitää herätä* why tomorrow must.3SG wake.up.INF *aikaisin?* early *'Why should Ø wake up early tomorrow?'*
Object NP	initial field	**Paperi** *pitää skannata* paper must.3SG scan.INF *'Ø has to scan the paper'*
	elsewhere	*Pitää ostaa* **kuulolaite** must.3SG buy.INF hearing.aid *'Ø has to buy a hearing aid'*
	no	
Past tense form	yes/no	**Piti** *lähteä pois* must.3SG leave.INF away *'Ø had to leave'*
Presence of discourse particles	yes/no	*Tuo pitäisi* **kyllä** *tietää* DEM must.3SG PTC know.INF *'Ø should surely know that'*
Discursive		
Position of the clause inside a turn	turn-initial turn-final neither (middle) both	See example (4)
Random effects		
Lemma of the infinitival complement		
Speaker		

---

8   The conditional *pitäisi* often translates best to *should*.

The morphosyntactic annotation in the Arkisyn database makes it possible to automatically retrieve all variables apart from the turn structure-related variables (on turn construction, see Sacks, Schegloff & Jefferson 1974). In the annotation of the corpus, the turns are organized directly in accordance with the transcriptions, which means that certain characteristics of conversation, such as overlapping, could cause a turn to be scattered into two or more annotational segments, although in reality, there would be an ongoing utterance. Therefore, all the turns were manually examined before extracting the information.[9]

4. Results and analysis

The results are introduced as follows: Section 4.1 presents the results of the GLMM and shows the general tendencies between the CovertSubj and SpeechActSubj utterances, whereas Section 4.2 focuses solely on their statistically significant differences, provides a qualitative analysis of the differences in talk-in-interaction, and puts these findings into context with respect to previous research. Sections 4.3 and 4.4 do the same for CovertSubj and 3PersSubj utterances.

4.1. First comparison (CovertSubj vs. SpeechActSubj): Model fit and general tendencies

The model is able to capture the observed variance only moderately ($R^2_{marginal} = 0.29$, $R^2_{conditional} = 0.38$).[10] The random effects contributed less than 10% to the amount of variance explained (for the measure of variance explained, see Winter 2013: 5), as indicated by the $R^2_{conditional}$. Altogether eight out of ten fixed effects and neither of the random effects were statistically significant (all but the turn-initial position of the clause and adverbial conjunction), according to the likelihood ratio test. Variance inflation factors (VIFs) were used to estimate collinearity between the variables. All the VIFs are well below 3, which indicates that collinearity was not an issue (Zuur, Ieno & Elphick 2010). The results are presented in Table 2.

9   The turn-coding in this study should be regarded only as indicative: Its intention is to do justice to the data, but differing views in the coding are also possible.

10 When interpreting the proportions of explained variance, it should be kept in mind that the studied constructions do not occur in a complementary distribution in the strict sense.

Table 2. Results of the likelihood ratio test between CovertSubj and SpeechAct-Subj. Statistically significant differences are marked with *.

Fixed effects	VIF	df	χ²	p-value	Favored by
Adverbial in the initial field	1.03	1	13.82	<0.001*	CovertSubj
Adverbial conjunction	1.02	1	0.11	0.737	CovertSubj
että-clause	1.11	1	4.06	0.044*	CovertSubj
Conditional mood	1.08	1	8.38	0.004*	CovertSubj
Interrogative form	1.11	1	11.38	<0.001*	SpeechActSubj
Object NP	1.07	2	17.37	<0.001*	Ambiguous
Past tense	1.04	1	7.28	0.007*	SpeechActSubj
Presence of discourse particles and interjections	1.06	1	4.23	0.040*	SpeechActSubj
Position of the clause inside a turn	1.06	3	6.76	0.080	Ambiguous
Random effects	**Standard deviation**[11]			**p-value**	
Lemma of the infinitival complement	0.411			0.574	
Speaker	0.548			0.142	

With regard to the semantico-grammatical and discursive characteristics, the CovertSubj utterances clearly differ from those of SpeechActSubj: Adverbials in the initial field, *että*-clauses, and conditional mood are all statistically significant and favored by CovertSubj utterances, whereas the presence of discourse particles and interjections, interrogation, and past tense forms are also statistically significant but instead favored by SpeechActSubj utterances. Object NP also proves to be statistically significant, but it is its location that separates the CovertSubj and SpeechActSubj utterances: In the former, the object NP is typically in the initial field, whereas in the latter, it is located elsewhere in the clause. Additionally, the position of the clause inside a turn is close to being statistically significant, and the one-clause turns (i.e., simultaneously turn-initial and turn-final) in particular make up most of the difference, as illustrated and explained more comprehensively in Section 4.2.

---

11   Logarithmic scale is used in Tables 2 and 3.

## 4.2. Differences between CovertSubj and SpeechActSubj

In this section, I will focus on the statistically significant differences between the two groups. First, I will examine the fixed effects favored by the CovertSubj utterances (in the order of their significance) and then move on to the fixed effects favored by the SpeechActSubj utterances. I will show that the interactional tasks related to planning and proposal-making account for the tendencies of CovertSubj compared with SpeechActSubj, whereas the latter are more oriented to the necessity and obligation itself.

Features favored by CovertSubj

The object NPs (in the initial field), adverbials in the initial field, conditional mood, and *että*-clauses are all statistically significant and favored by the CovertSubj utterances. In the context of Finnish zero-person constructions, both adverbials and object NPs bear a similar function as the preverbal theme, as they are used to define the conditions that affect the implied referent (Laitinen 2006: 214–215), whereas the conditional mood is commonly present in contexts of proposal-making and planning (e.g. Varjo & Suomalainen 2018, Varjo 2019; cf. the Estonian personless conditional, Pajusalu 2015: 47).

Object NP

Object NP, along with adverbial in the initial field, proved to be the most significant fixed effect, $\chi^2(2) = 17.37$, $p < 0.001$, yet it is still ambiguous, as Figure 1 shows: Object NPs show a tendency toward CovertSubj utterances in the initial field, but the same tendency is not present elsewhere in the clause. In other words, apart from the object NPs in the initial field, object NPs are more common among Speech-ActSubj utterances.
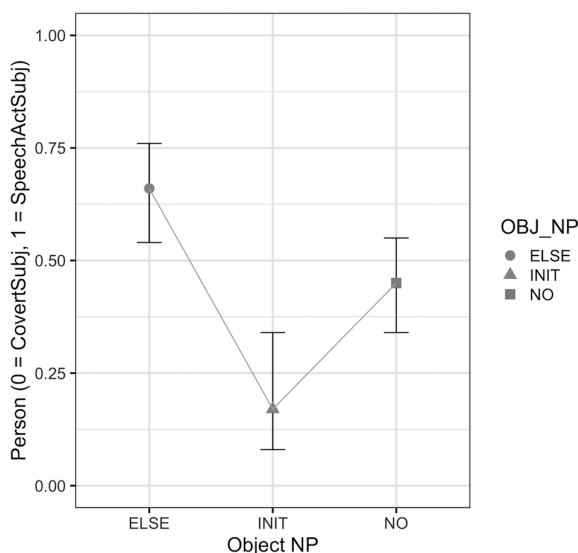
Figure 1. The estimated probability of person being SpeechActSubj vs. Covert-Subj as a function of Object NP.

The object NPs in the initial field of zero-person constructions are often pronominal, especially demonstratives (Varjo 2019), which indicates that they are at the center of attention (see also Gundel, Hedberg & Zacharski 1993 on Givenness Hierarchy).

Zero-person constructions with an object NP in the initial field typically function as proposals or directives concerning the referent of the object NP in the theme position (e.g. Varjo 2019). This is also true with the CovertSubj utterance in example (2),[12] where Heikki in line 7 makes an argument for emptying the walk-in closet because it is so full of stuff.

(2)   [SG444]

    1 Heikki: *nn [jos sen saa seinää   ] vasten pystyy.*
                'If one is able to get it in an upright position toward the wall'

    2 Teppo: *[sit ku e̲i pe̲laa ni (sit,)] (.)*
                'Then when one does not play'

    3 Teppo: *m[m.]*
                'Yeah'

_____

12   Transcription conventions are explained in an appendix.

 4 Tuomas: *[ni]in no joo.* (0.2) *#no joo.#* (0.2)
'Well, yeah, yeah'

 5 Heikki: *meil a-* (.) *meille mahtuis iha hyvi siihe vaatehuoneesee.* (0.2)
paitsi
'We a- we would have good room for it in the walk-in closet
except'

 6 *et se on täynnä kaik↑kee romuu nytte ni,* (0.4)
'it's filled with all sorts of junk now'

 7 **pitäs   tyhjentää.**
must.3SG.COND   empty.INF
'Ø should empty (it)'

 8 **se   pitäs   kyllä   <muutenkin>**
DEM   must.3SG.COND   PTC   anyway

 9 **tyhjentää.** (1.2)
empty.INF
'Ø should empty it anyway'

10 Tuomas: *(°mut jos°) sä tarviit niit °joskus vielä.°* (.)
'But if you need them some day'

11 Heikki: *heh heh heh* ((Heikki shakes his head and drinks out of
a cup)) (3.6)

In example (2), participants are in the middle of playing the Dungeons and Dragons board game, and right before the excerpt, they have discussed building a game board of their own. The only problem is that it is so big that it is rather difficult to store unless one makes it foldable. Heikki in line 1 admits that it would be possible to store the board if one could just lean it against the wall. In line 5, Heikki continues by saying that the board could fit into their walk-in closet if it were only emptied of all the junk first. Heikki then uses a CovertSubj utterance in conditional mood, first focusing the attention on the action in line 7 and then on the walk-in closet itself in lines 8–9 with an object NP in the initial field, which could be interpreted as an implied proposal – perhaps trying to engage his brother Ville, with whom he shares the flat and who is also present in the conversation, in joint planning (cf. Couper-Kuhlen & Etelämäki 2015: 12, Pajusalu 2015: 46–50 on promoting [joint] future activities with open personal constructions). Ville, however, stays silent and does not respond to his brother's comment in any way, although

there is a rather lengthy pause after Heikki's turn. Tuomas then takes the floor, mumbling to Heikki. Heikki's response in line 11 is rather laconic and leads to an even longer pause, possibly implying he did not get the response he was hoping for.

Adverbial in the initial field

The use of adverbials in the initial field (along with the object NPs) proved to be the most significant difference between CovertSubj and SpeechActSubj utterances, $\chi^2(1) = 13.82$, p < 0.001, as Figure 2 also depicts.
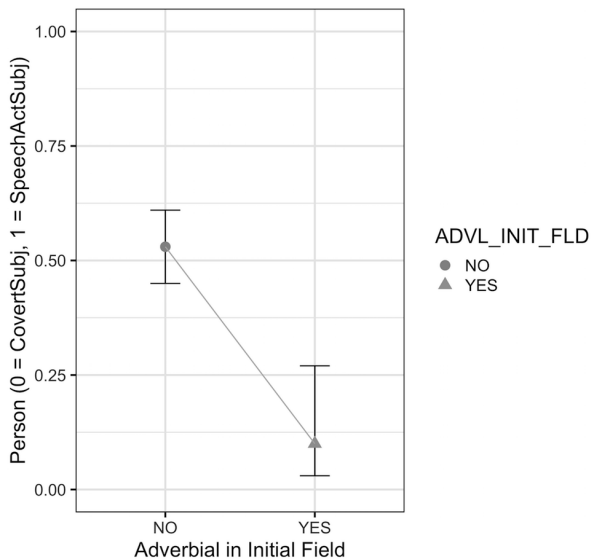
Figure 2. The estimated probability of person being SpeechActSubj vs. Covert-Subj as a function of Adverbial in the initial field.

Adverbials, especially in the initial field of zero-person constructions, typically form necessary or sufficient conditions for the action, process, or other event described in the construction,[13] for instance, *Täällä ei opi mitään* 'You don't learn anything here' (example and

---

13   Especially with concrete verbs expressing a change of state such as *Täällä jäätyy* 'One freezes here', the zero-person construction is typically interpreted so that the event is inevitable/automatic in those conditions (e.g. Hakulinen et al. 2004: §1353).

its translation from Hakulinen & Karttunen 1973). Therefore, it is not surprising that adverbials in the initial field of the clause are far more common among CovertSubj utterances. What this therefore suggests is that the CovertSubj utterances carry some typical characteristics of zero-person constructions (e.g. Hakulinen et al. 2004: §1353) when it comes to their information structure (Hakulinen et al. 2004: § 1370). In this respect, they resemble the object NPs in CovertSubj utterances described in the previous section. An example of and adverbial in the initial field can be seen in example (5).

Conditional mood

The CovertSubj utterances favor the conditional mood in a statistically significant fashion, $\chi^2(1) = 8.38$, $p = 0.004$, as Figure 3 depicts. Markers for creating a hypothetical or otherwise nonfactual context are commonly used in zero-person constructions, as they typically create an open reference. Besides the conditional mood, the conditional *if–then* frame bears a similar function (in this study portrayed by the factor of adverbial conjunctions).
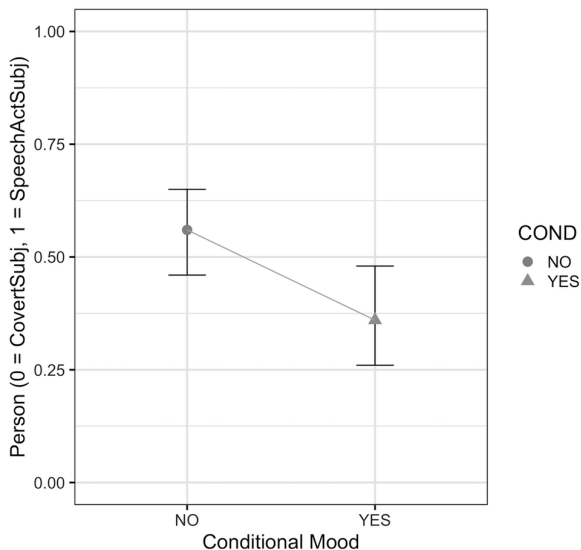
Figure 3. The estimated probability of person being SpeechActSubj vs. CovertSubj as a function of Conditional mood.

Zero-person constructions have a strong tendency toward the conditional mood (e.g. Varjo & Suomalainen 2018: 347). Zero-person constructions with their predicate verb in conditional mood are particularly used in making a proposal or expressing personal wishes (Varjo 2019; see also Lindström et al. 2016). This is also true when it comes to the CovertSubj utterances. Example (3) shows a CovertSubj utterance that has its predicate verb in conditional mood occurring in a proposal-making context, as Mika expresses his interest to go and see a movie called *Donnie Brasco*.

(3)   [SG121]

1 Mika:	*tää*	*tää*	*tää (.)*		
	DEM	DEM	DEM		

2	*tää*	*Donnie*	*Brasco*	*pitäs*	*käyä*
	DEM	Donnie	Brasco	must.3SG.COND	go.INF

3   *kattoo,*
watch.INF.INE
'This this this this Donnie Brasco Ø would have to go and see'

4 Jere: *mm:y,*
'Yeah'

5 Tero: *nyt meni (.) vähä [vituiks] tää homma,*
'Now this thing went south'

6 Mika: *[se   ]*
'It'

7 Mika: *Al Pacino ja Johnn[y Depp.   ]*
'Al Pacino and Johnny Depp'

8 Jere: *[se kiinnostas kyllä,]*
'That would be interesting'

In lines 2–3, Mika uses a CovertSubj utterance to plan a future activity, namely addressing that he is interested in seeing the movie (*tää Donnie Brasco* being the object NP of the clause). Jere immediately aligns by producing a discourse particle *mm* (e.g. Sorjonen 2001: 36, 95), whereas Tero is still engaged in other activities. Mika gives further reasoning as to why he is interested in the movie in line 7 in the form of a free NP (see Helasvuo 2019), naming two actors who have roles in the

movie. Jere again in line 8 aligns with what Mika has said, this time by producing a longer response. By using a referentially open NEC, Mika is testing the water, trying to get others to express their opinions on going to see the movie. At the same time, Mika leaves it open to go just by himself in case no one else is interested – an example of expressing one's personal wishes (cf. Varjo 2019: 71, Pajusalu 2016: 46–47). By contrast to example (2) in which the topic was mutually accessible, the CovertSubj utterance in (3) refers to the speaker's personal sphere.

Position of the clause inside a turn

The position of the clause inside a turn did not prove to be statistically significant overall, $\chi^2(3) = 6.76$, $p = 0.080$, but as depicted in Figure 4, clauses that are simultaneously turn-initial and turn-final still have a strong tendency toward the CovertSubj utterances. Interrogative forms, on the other hand, have a tendency toward SpeechActSubj utterances. Therefore, it can be concluded that simultaneously turn-initial and turn-final CovertSubj utterances cannot be explained by utterances that act as the first turn of question–answer adjacency pairs.
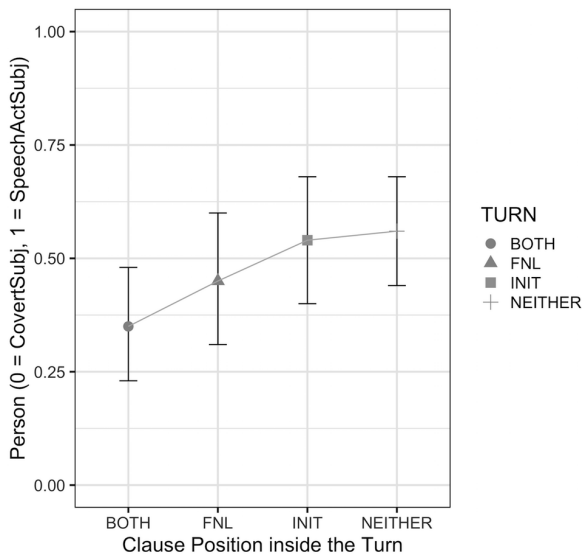
Figure 4. The estimated probability of person being SpeechActSubj vs. CovertSubj as a function of Position of the clause inside a turn.

Interestingly, CovertSubj utterances still seem to provide a transition-relevance place (see Sacks, Schegloff & Jefferson 1974) more often than SpeechActSubj utterances. This could be due to their use in, for instance, in promoting (joint) action and planning or in giving advice or issuing directives (e.g. Couper-Kuhlen & Etelämäki 2015), or as reactive utterances for taking a stance, in which case they are also often turn-initial (see Varjo 2019). An example of the use of a CovertSubj utterance in both a turn-initial and turn-final position (i.e. constituting a turn by itself) is provided in example (4), in which Niko in line 3 reacts to what has been said previously about a girl linked to his brother Iikka (the nickname of Iiro) by using a CovertSubj utterance (cf. Varjo 2019: 74–77 on reactive zero-person constructions with modal verbs that do not express necessity).

(4) [SG441]
    1 Elli: *s:e on Iikan ka:veri ei se mikää tyttöystävä [oo].*
              'She is a friend of Iikka, not a girlfriend'

    2 Iiro: *[hh]ho* (0.3)

3 Niko:	**no**	**kylhä**	**sitä**	**nyt**	**pitää**	**vähä**
	well	PTC.CL	DEM.PART	now	must.3SG	a.little

4	***kiusatas***	***£siitä£.*** (1.5)				
	tease.INF	DEM.ELA				

              'Well certainly Ø has to tease him a bit about that'

    5 Iiro: *siis Virvallaha on niinku jo po#i-*
              'Virva does already like have a bo-'

    6   *V-# Virvallaha o iham poikaystävä ja kaikkee n- et siis niinku,*
              'Virva does have a boyfriend and all so that like'

The CovertSubj utterance in lines 3–4 includes plenty of discourse particles, such as the particles *no* 'well' and *kyllä* 'certainly', along with a clitic, *-hän*, which marks the claim as something that should be obvious (e.g. Hakulinen et al. 2004: §830–§832). Niko uses the CovertSubj utterance to take a stance on the ongoing discussion about a girl with whom Iikka is spending a lot of time but who still has a boyfriend other than Iikka. Although the girl is not Iikka's girlfriend, Niko presents the situation as providing motivation for teasing: The reference of the CovertSubj utterance is left open, but Niko uses it to entitle his own actions.

että-clause

The *että*-clause proved to be statistically significant in the comparison, $\chi^2(1) = 4.06$, p = 0.044, and the difference between CovertSubj and SpeechActSubj is depicted in Figure 5.
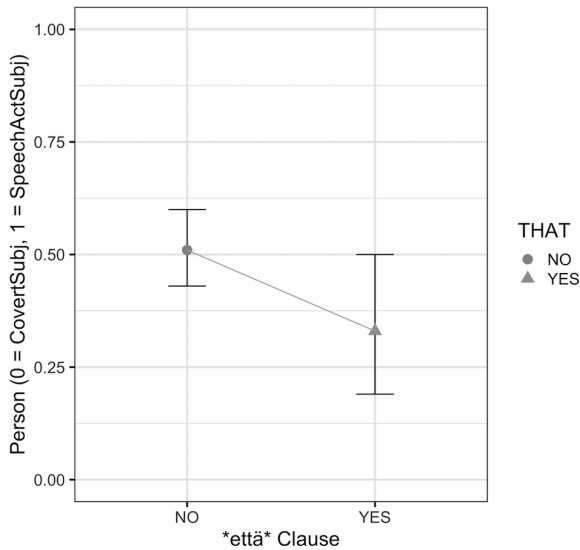
Figure 5. The estimated probability of person being SpeechActSubj vs. Covert-Subj as a function of *että*-clause.

Particularly in spoken interaction, the use of *että* differs from its more formal use: Although it has traditionally been considered a subordinating conjunction or a complementizer, in conversational Finnish, its main tasks are instead related to metacommunicatively organizing the participant framework (see Koivisto, Laury & Seppänen 2011). Koivisto et al. (2011) argue that even if *että* occurs in a complement-taking predicate phrase, it is not syntactically a subordinator, but it typically projects that there is more to come (and thus regulates turn-taking). Additionally, the utterance-initial particle is used to organize the participant framework by indexing what is to follow as a paraphrase, a candidate understanding, or an upshot of what has already been said (Koivisto, Laury & Seppänen 2011: 71, 96). Zero person therefore seems well suited to turns in which *että* is used in such metacommunicative functions, for

example in contexts of planning as in (5), as the participants discuss travelling back home the following morning.

(5)   [Sapu119]

   1 Jatta: *mihi aikaa te meinaatte aamust lähtee?*
             'At what time are you planning to leave in the morning?'

   2 Netta: *noh.*
             'Well'

   3 Alina: *hh varmaa yheksä, kymmene.*
             'Probably something like nine, ten'

   4 Netta: *nii, jotai sellast.*
             'Yeah, something like that'

   5 Alina: **et**   **ko**   **yhelttoist**   **viimeistää**   **pitää**
             COMP   when   eleven.ELA   at.latest   must.3SG

   6   **>ainaki<**   **bussil**   **lähtee**
             at.least   bus.ADE   leave.INF
             'Because Ø has to leave by bus at eleven at the latest'

   7   *ja sit me käydää ennen tääl jossai syämäs.*
             'and then we'll go and eat here somewhere before (that)'

   8 Anni: *m,*
             (2.0)

   Jatta is the one providing accommodation for her little sister Netta and her friends for a festival weekend. In line 1, she asks at what time her guests plan to leave the following morning. Alina first produces a time estimate in line 3, and Netta aligns with this in line 4 by producing particle *nii* (see Sorjonen 2001: 139–140, 154). Then, in lines 5–6, Alina continues and uses an *että*-clause (that is also a CovertSubj utterance) to provide a conclusion/reasoning for their schedule by saying that they need to take a bus that leaves at eleven o'clock at the latest, and before that, they will need some time to eat. In Alina's turn, *että* acts as the first element in a syntactically and prosodically independent unit and is used to comment on what is being said (cf. Koivisto, Laury & Seppänen 2011: 82). As such, *että*-initiated turn both projects backwards to Jatta's question in line 1 and also has a forward-projecting function

in a planning sequence (cf. Koivisto, Laury & Seppänen 2011: 85), as Alina takes an active role in implementing the action at a given time, an example how *että* has the function of regulating the participant roles in interaction (Koivisto, Laury & Seppänen 2011: 86). This interpretation is reinforced by the clausal connector *kun* 'as, when' which is here used to imply that something important is to come (cf. Sorjonen 2001: 227). Alina's CovertSubj utterance is a typical zero-person construction in that it has an adverbial in the initial field framing the conditions that affect the implied referent (Laitinen 2006: 214–215) – in this case, anyone in their group.

Features favored by SpeechActSubj

The interrogative form, the past tense, and the presence of discourse particles were all statistically significant and favored by the SpeechActSubj utterances. Some of these features come as no surprise: For example, the past tense forms – especially the simple past – are typically used in narrative sequences in Finnish, and it is therefore natural that they typically have a specific reference, whereas in generic expressions, present tense is usually the preferred tense (e.g. Jokela & Plado 2015: 91). Some differences, on the other hand, are more complex, such as the presence of discourse particles and interjections, as the discourse particles in particular have a central role in constructing the action performed by the turn (e.g. Hakulinen 2001).

Past tense forms

As Figure 6 shows, past tense forms demonstrate a clear tendency toward SpeechActSubj, $\chi^2(1) = 7.28$, $p = 0.007$. In this study, the three past tense forms in Finnish, namely the simple past, present perfect, and past perfect, have been conflated to avoid data sparsity, but the simple past is still by far the most common of the three (with the conditional mood, only the present and the present perfect are possible).
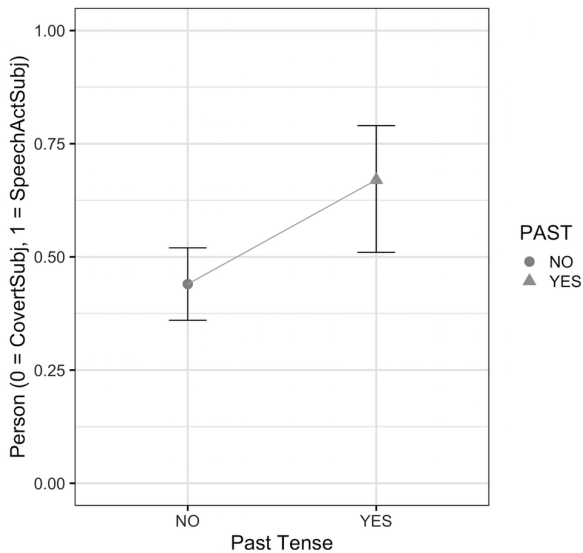
Figure 6. The estimated probability of person being SpeechActSubj vs. Covert-Subj as a function of Past tense.

Zero person quite often receives a more specific interpretation in past tense contexts (Laitinen 2006: 212). Tense is also firmly entwined in verb semantics through narrativity: Narratives often involve verbs of action and motion, and as there are fewer intervening clauses, the protagonist remains clear based on the context (Lindström et al. 2009). Example (6) shows a personal narrative where Jaana uses both a CovertSubj utterance and a SpeechAct-Subj utterance referring to herself when she is talking about her history with psoriasis.

(6) [SG437]

1	Tuula:	↑*sillonha* **mulle** *p<u>u</u>hkes se ps<u>o</u>riaassis?*
		'At that time I broke out in psoriasis'
2		*.mthhh ja sitten;* (.) *sillon ku* **mä** *olin o-*
		'and then when I had b-'
3		*viis vuotta ollu. .hhhh* ↑*k<u>o</u>tona lasten kanssa;* (.)
		'been at home with the children for five years'

4	***ja, (.)***	***>piti***	***lähteet<***	***takasi***	***töihin.***
	and	must.3SG	leave.INF	back	work.PL.ILL

'and Ø had to leave back to work'

5   *.hh ni sillon;* (0.4) *mt sem-,* (0.2) *puhkes pääs- päänahassa*
'right then it broke out in the scalp'

6   *.mthh että #e# ke- ↑kesällä;* (0.2)
'like in the summer'

7   ***mä*** *olin menos syksyllä töihin;*
'I was going back to work in the autumn'

8   *.hh nin kesällä.* (0.6) *mt alko se hilse nouseen sinne.* (.) *[et]*
'so in the summer the dandruff started to break out there so'

9 Jaana: *[joo.]*
                                                                  'Yeah'

10 Tuula: *koko pää °>oli ku<° päänahka oli niinkun i↑han yhtä;* (0.4)
yhtenäistä laattaa,
'The whole head was like, the scalp was like one solid slab'

11 Jaana: *[joo.]*
'Yeah'

12 Tuula: *[.hhh] ja sitten .hh se edellinen yöh;* (.)
'and then the previous night'

13	***kum***	***mum***	***piti***	***mennä***	***>seuraavan***
	when	1SG.GEN	must.3SG	go.INF	next.ESS

14	***päivänä<***	***töihin***
	day.ESS	work.PL.ILL

'when I had to go back to work the next day'

15   *..thhh niin sit se rupes mär°kimään°.* (1.8)
'and just then it began to fester'

The CovertSubj utterance in line 4 is a coordinate clause and preceded by a first-person pronominal subject (cf. Hakulinen et al. 2004: §1348) in line 2. There are also several other first-person pronouns and verbal subject markers in the sequence. This seems to be typical of other CovertSubj utterances in personal narratives in the data as well. In lines 13–14, however, Tuula uses a SpeechActSubj utterance. Unlike in line 4, there are no first-person markers close by, as the nearest one is at a distance of three clauses in line 7 (cf. Travis 2005: 344). Both NECs

express external necessity, that is to say, refer to circumstances external to the participant, and include verbs of movement.

Another common pattern in the data with a past tense form and an explicit subject (often in interrogative form) has a metacommunictive function of projecting that something more is to follow, as in example (7), in which Iina is giving her mother-in-law news about her son Tomi and his old friends.

(7) [SG446]

   1 Iina: *mialummin ottaa sit sii̲hen lainaa*
            'One rather takes a loan for that then'

   2   *ku johonki tämmösee kerrostaloko[p°pii et°,]*
            'than for a small flat like this'

   3 Ritva: [<a̲iv]an> h #joo#.
                               'Yeah indeed'

   4 Iina: *ka   n> h   °sitte°,*
            'then'

   5   (0.3)

   6 Ritva: *se=on sit sen °ajan murhe°;*
            'That is nothing to worry about now'

   7 Iina: *nii[npä. h]*
            'Exactly'

   8 Ritva: *[mm hi] hi (.)*
            'Yeah'

   9   *°et° kuha aika [rientää.   ]*
            'as long as time flies'

  10 Iina: **[°mitä   mun   pi]ti   sanoo   viäl°**
                  what.PART 1SG.GEN must.3SG say.INF   still
                  'What else was I supposed to say now'

  11   *nii̲ joo; (.) toi (0.3) ny̲thän noi kaikki täyttää kolkyt tosiaan tänä vuonna noi*
       'Oh right, this year they all are going to turn 30'
       ((Iina smiling))

  12   *kaikki [Tomin] kaverit,*
       'All of Tomi's friends'

In line 10, Iina uses a SpeechActSubj utterance with *pitää* in the past tense, but this time as a strategy to give her time to think about what she is supposed to say next, namely that Tomi and his friends all turn 30 soon.

Interrogative form

As indicated by Figure 7, interrogation is one of the four variables that have a significant tendency toward SpeechActSubj utterances, $\chi^2(1) = 11.38$, $p < 0.001$. In this study, interrogation was encoded based on the presence of interrogative pronouns and clitics, which are also the most common ways to mark a clause as a question in Finnish.
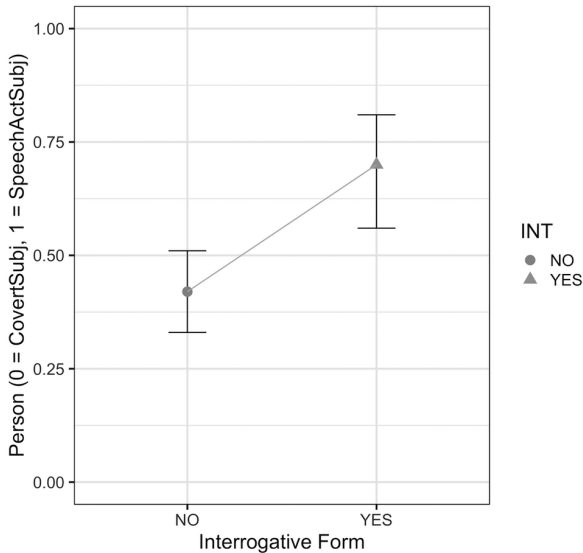
Figure 7. The estimated probability of person being SpeechActSubj vs. Covert-Subj as a function of Interrogative form.

As questions are addressed by the speaker to the recipient in exchanges of information or requests etc., it is not surprising that the results here show a preference toward speech act pronouns as subjects. However, even the referentially specific interrogative NECs seem to be used to initiate action rather than asking for information, as in example (8), in which Lotta, Milja, and Oona are doing their homework exercises together.

(8)   [SG120]

　　1 Lotta:　*haluutsie kattoo täst jotai.*
　　　　　　'Would you like to see something from here'

　　2 Oona:　*ei ku mie nään kyllä.=*
　　　　　　'No, I (can) see well'

　　3 Milja:　*=tota, (0.5)*
　　　　　　'Umm'

　　4 Milja:　**mm. pitäskö**　　　　**meiä**　　　**suunnitella**　　**sitä**
　　　　　　must.3SG.COND.Q　　WE.GEN　　plan.INF　　　DEM.PART

　　5　　　　**retkee(ki).h**
　　　　　　trip.PART.CL
　　　　　　'Yeah, should we plan that trip as well'

　　6 Lotta:　*mm-m,*
　　　　　　'Yeah'

　　7 Oona:　**kyl**　**kai**　　**meiä**　　**pitäs.**
　　　　　　PTC　I.guess　WE.GEN　must.3SG.COND
　　　　　　'I guess we should'

　　　In the middle of comparing their calculations, Milja uses a referentially specific NEC with a first-person plural pronominal subject to shift the focus to a trip they should be planning while they are all together. Lotta shows alignment with a discourse particle *mm* (see Sorjonen 2001: 36, 95), whereas Oona uses the same structure of plural first-person pronominal subject and *pitää* in conditional mood, what could constitute an example of syntactic priming (e.g. Travis 2005), yet adding a modal particle chain *kyllä kai* (Hakulinen et al. 2004: §1603) at the beginning of the utterance to express a slight reluctance or hesitation. Compared with the referentially open interrogative NEC in example (9), which is used to propose an action, the emphasis in (8) is more clearly on the interlocutors' external need to perform the activity.

(9)   [SG444]

　　1 Tuomas:　hhh (0.6) #joo# ompa ↓hohho↑hoo ↓hoo ↑hylmää.
　　　　　　　'Yeah it is cold all right'

　　2　　　　　((Tuomas holding his hand in front of his mouth))

　　3　　　　　(.)

4 Tuomas: m,

5   (1.8)

6 SOMEBODY: [kylmää?]
               'Cold'

7 SOMEBODY: >[(joo)]<
               'Yeah'

8 Tuomas: o. (0.4)
               'It is'

9 Heikki: **>p'täiskö<**   **ottaa**   **>se<**   **limupullo**
             must.3.SG.COND.Q   take.INF   DEM   bottle.of.soda

10   **#vaan#**   **tähä**   **pöytään**
             just   DEM.ILL   table.ILL
             'Should Ø just take that bottle of soda here to the table?'

11 Teppo: >kyllä se< sieltä menee. (1.8)
   'It'll be used up from there'

When zero person is used in interrogative contexts, it often provides a more suitable alternative as a generic question than, for instance, the open second-person singular form (Laitinen 2006: 221–222). In the excerpt, Teppo, Ville, Tuomas, and Heikki are having some soda in the middle of playing Dungeons and Dragons. Tuomas comments in line 1 that the soda is very cold. This, after a pause, elicits a response from Heikki, who in lines 9–10 proposes taking the whole bottle out of the fridge and to the table by using an interrogative CovertSubj utterance. The question is a general one in that it does not specify who should be the one to get the bottle. This can be seen also in Teppo's response in line 11 in which there is no personal reference, only a very general statement to take an opposing stance in the matter.

Presence of discourse particles and interjections

As Figure 8 shows, discourse particles and interjections[14] favor SpeechActSubj utterances, $\chi^2(1) = 4.23$, p = 0.040. Discourse particles

---

14 These include clitics -*hAn*, -*pA*, -*kin*, and -*kAAn*; particles *kyllä*, *tosiaan*, *kai*, *ai*, and *no*; and all the interjections (according to the Arkisyn annotation).

are an inherent part of conversation dynamics. They are inter-subjective and used to organize the social interaction in different ways: For example, *nyt* 'now' is used in exclamations, understatements, exhortations, and requests (Hakulinen 1998), whereas *kyllä* 'sure' often conveys reaction to something and is a way to deal with delicate actions or differing opinions (Hakulinen 2001). Regarding non-necessive zero-person constructions in conversational Finnish, discourse particles are present especially in constructions that are either in turn-initial or turn-final position (or both). Furthermore, they show a tendency to co-occur with modal verbs and a clause-initial object NP and are typically used in expressing personal stance or giving directives (Varjo 2019: 74–77).
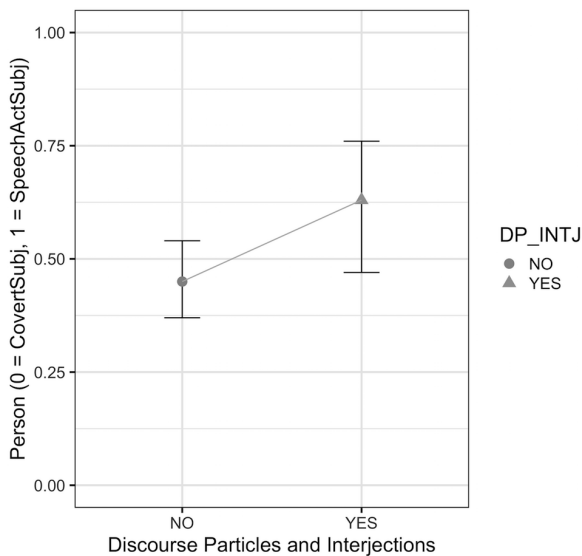
Figure 8. The estimated probability of person being SpeechActSubj vs. Covert-Subj as a function of Discourse particles and interjections.

In example (10), the family members discuss Niko, whose military service is about to end sooner than expected. In line 6, Niko's mother Elli uses a referentially specific NEC to emphasize the need for Niko to remember to terminate his lease in good time before the end of his military service.

(10) [SG441]

  1 Elli: *nii mut siis,* (0.4) *onk sulle nyt luvattu*
               'Yeah but have you already been promised'

  2   *e[t s]ä pääset tammikuussa?*
               'that you get out in January'

  3 JOKU: [.nf]

  4 Niko: *no: on se aika varma.*
               'Well it is quite certain'

  5 Mari: *[o:ho?]*
               'Oh'

  6 Elli:

[no	***sun]***	***pitää***	***sitte***	***se***	***kämppä***
PTC	2SG.GEN	must.3SG	then	DEM	flat

  7

irtisanoo	***sillee***
terminate.INF	like

  8

et	***se***	***pitäs***	***periaat[tees***	***ir]tisanoo***	***sit***
COMP	DEM	must.3SG	in.principle	terminate	then

  9

jo	***marraskuun***	***aikana.***
already	November.GEN	during

'Well you must then terminate the lease so that Ø should in
principle terminate it already in November'

  10 Mari: *[nyt jo,]*
               'Already'

  11 Niko: *[mm-m,]*
               'Yeah'

Before the excerpt, Elli has already emphasized the importance of sorting out the housing issue during the ongoing weekend. In line 1, she returns to the topic, making sure that Niko will get out of his service in January. In Niko's response in line 4, he claims it is quite certain, to which Elli responds with a SpeechActSubj utterance beginning with *no* 'well'. Elli's turn itself is responsive and aims to take a stance on what Niko should do next (cf. Vepsäläinen 2019: 65–67). The emphasis is on the external necessity, as in (5): Elli gives a directive to her son. Interestingly, after the SpeechActSubj utterance, Elli switches to a CovertSubj utterance with a clause-initial object NP in lines 8–9 as

she formulates a generalization about how anyone in the same circumstances should act, thus ending her turn.

4.3. Model fit and general tendencies in the second comparison

The second model is able to capture the observed variance only moderately, yet the random effects improved the observed variance remarkably ($R^2_{marginal} = 0.264$, $R^2_{conditional} = 0.788$). The variance of the random variables shows that this is almost exclusively due to the lemma of the infinitival complement, whereas the variance across speakers has virtually no effect. Out of the fixed effects, only the past tense was a statistically significant predictor, but the object NP and interrogative form also came close to the 0.05 level. Similar to the first comparison, VIFs were used to estimate the collinearity between the variables. All the VIFs were well below 3 also in the second comparison. The results are presented in Table 3.

Table 3. Results of likelihood ratio test between CovertSubj and 3PersSubj. Statistically significant differences are marked with *.

| Fixed effects	VIF	df	χ^2	p-value	Favored by
Adverbial in the initial field	1.03	1	2.03	0.154	CovertSubj
Adverbial conjunction	1.05	1	0.64	0.424	CovertSubj
että-clause	1.16	1	0.75	0.386	CovertSubj
Conditional mood	1.09	1	0.00	0.967	CovertSubj
Interrogative form	1.11	1	3.32	0.068	CovertSubj
Object NP	1.14	2	5.00	0.082	Ambiguous
Past tense	1.38	1	7.67	0.006*	3PersSubj
Presence of discourse particles and interjections	1.18	1	0.79	0.374	
Position of the clause inside a turn	1.12	3	2.99	0.393	Ambiguous
Random effects	**Standard deviation**				
Lemma of the infinitival complement	2.851			<0.001*	
Speaker	0.000			0.999	

4.4. Differences between CovertSubj and 3PersSubj

In this section, I will focus on the statistically significant differences between the CovertSubj and 3PersSubj utterances. Because there are fewer differences between these two groups than there were in Section 4.2 with CovertSubj and SpeechActSubj utterances, I will start with the past tense forms, which show the only statistically significant difference (and favored by the 3PersSubj utterances), and then move on to an overview of the random effects, which improved the observed variance remarkably in this second comparison.

Past tense forms

The only statistically significant difference, and at the same time a commonality between SpeechActSubj and 3PersSubj, is the past tense forms that are significantly skewed toward both the referentially specific groups of NECs as opposed to CovertSubj, $\chi^2(1) = 7.67$, $p = 0.006$. With respect to the 3PersSubj utterances, this is depicted in Figure 9. As mentioned already in Subsection 4.2, narrative sequences as in (5) could explain the difference.
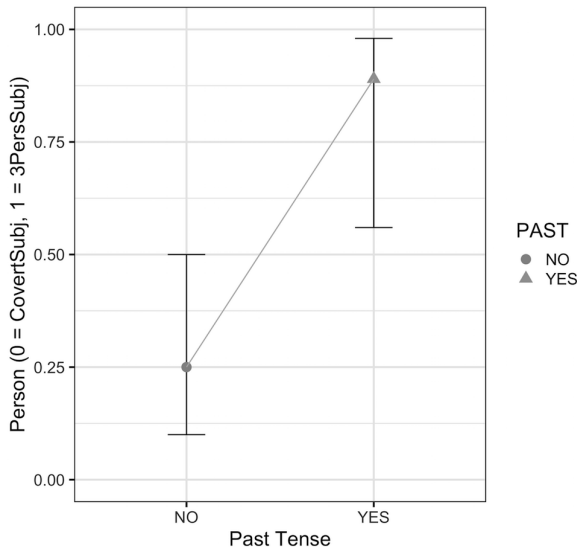
Figure 9. The estimated probability of person being 3PersSubj vs. CovertSubj as a function of Past tense.

Random effects

Among 3PersSubj, both random effects can be considered relevant, although only the lemma of the infinitival complement is statistically significant, $\chi^2(1) = 34.35$, $p < 0.001$. The random effects also add to the goodness of fit of the model remarkably (the proportion of explained variance with fixed effects only is 22% and with both fixed and random effects 68%). Although there are fewer differences among the fixed effects between CovertSubj and 3PersSubj than between CovertSubj and SpeechActSubj, the verb distribution in the infinitival complement reveals something essential about the different interactional functions the CovertSubj and 3PersSubj utterances have. Table 4 presents the most common infinitival complement lemmas in both groups (those that exceed 3% of all the occurrences).

Table 4. The lemma distribution of infinitival complements in all the groups.

CovertSubj	%	SpeechActSubj	%	3PersSubj	%
no infinitival complement	12.9	no infinitival complement	14.2	*olla* 'to be'	33.7
laittaa 'to put'	5.6	*mennä* 'to go'	6.3	*tulla* 'to come'	4.2
ottaa 'to take'	5.1	*sanoa* 'to say'	4.7	*mennä* 'to go'	4.2
tehdä 'to do/make'	4.5	*lähteä* 'to leave'	4.7	*käydä* 'to visit'	4.2
olla 'to be'	4.2	*käydä* 'to visit'	4.7	*saada* 'to get'	3.2
mennä 'to go'	3.9	*tehdä* 'to do/make'	4.2		
lähteä 'to leave'	3.7	*saada* 'to get'	3.7		
		ottaa 'to take'	3.2		

The differences in the lemma distribution among infinitival complements indicate that both CovertSubj and SpeechActSubj utterances are quite often used in turns that, for instance, are intercepted or occur as part of an adjacency pair (such as the answer in line 6 of example 8) and thus have no infinitival complement at all, whereas 3PersSubj utterances are quite conventionalized in their context of use, which can be seen as scarcity in the verb list: Five of the most frequent verbs cover almost 50% the utterances in 3PersSubj.[15] According to Jaakola (2004: 259),

---

15   The animacy of the third-person subject is not considered in this study.

necessity constructions without an infinitival complement are possible because the main function of necessity constructions is to emphasize that the subject is experiencing or undergoing an event, whereas the quality of the verb used as the infinitival complement greatly affects how the referent of the subject NP is seen in terms of agentivity and intentionality.

Example (11) demonstrates the use of a 3PersSubj utterance with *olla* as the infinitive complement as Jaana is engaged in telling a story about her old workplace.

(11) [SG438]

| 1 Jaana: | *aattele et ku, .hhhhmh ku tota me oltii siin tiskil ihan koko päivä* |
| | 'Think that we were there at the counter all day' |

| 2 | *siit ei voinu oikeen ↑liikkuu mihinkään* |
| | 'Ø could not really go anywhere' |

| 3 | *ku, .hhhh sitä jengii tuli ja meni.* |
| | 'when people kept coming and going' |

| 4 | (0.4) *no aattele >nyt et ku< kuus#sataa ihmist oli siellä#.* |
| | 'well think that there were six hundred people' |

| 5 | (.) *ni eihän sii[t vo]inu lähtee* |
| | 'so Ø surely could not leave' |

| 6 Tuula: | *[mm,]* |
| | 'Yeah' |

7 Jaana:	***et jonku piti aina siin olla***
	COMP someone.GEN must.3SG always there be.INF
	'Someone hade to stay there at all times'

In lines 2 and 4, Jaana uses a modal yet non-necessive zero-person construction to make a generalization about her own experiences (see Varjo 2019: 76–78) and to describe the situation to her sister. Finally, in line 7, Jaana uses an *että*-clause with a third-person subject to sum up what she has just said about working at the reception of a theatre. This instance is one of the few in the data set in which the referent is not only animate but also human, and the *olla* complement is used as a static location verb and not a copula.

5.   Conclusion

To provide an answer to the question of how the referentially open necessity constructions differ from the referentially specific ones by their interactional tasks and their contexts of use, the quantitative analysis revealed notable differences, especially in the first comparison between the CovertSubj and SpeechActSubj utterances, as altogether seven out of nine fixed effects proved to be statistically significant. The qualitative analysis shows that these differences in the semantico-grammatical and discursive features appear on the pragmatic level: CovertSubj utterances tend to be used in proposals and directives (e.g. object NPs and adverbials in the initial field, example 2), planning, suggestions and expressing personal wishes (e.g. conditional mood, example 3), reacting to something that is discussed (position of the clause inside the turn, example 4), and organizing the participant framework, especially while implementing an action (*että*-clause, example 5). SpeechActSubj utterances, on the contrary, are more oriented toward emphasizing the necessity or obligation of taking an action rather than directly implementing it (e.g. interrogative forms, example 8) while also expressing an affective stance (presence of discourse particles and interjections, example 10).

A comparison between CovertSubj and 3PersSubj further underlined that each of the three groups has its own unique characteristics. Although there were not as many statistically significant differences among the fixed effects in the second comparison, the lemma distribution of infinitival complements together with the tendency for the past tense forms shows a fundamentally different profile compared with CovertSubj and even with SpeechActSubj: The 3PersSubj utterances are frequently used in the meaning of 'must/should be' or in narrative sequences, as *olla* 'to be' and the common verbs of activity and movement make up approximately 50% of all the infinitival complements.

The distribution of NECs shows that the CovertSubj utterances are far more common in the data set than SpeechActSubj or 3PersSubj utterances. This indicates that as an interactional resource, the NECs are, above all, connected with matters that are interpersonally shared among the participants, and the reference is negotiated in the context.

On a more general level, this study contributes to interactional linguistics and the study of spoken interaction by combining quantitative methods with the study of personal reference and necessity

constructions. It provides an example of how GLMM can help in analyzing the different aspects of necessity constructions and shows the advantages of utilizing a well-established statistical method together with a morphologically and syntactically annotated corpus of spoken interaction when tracing the usage-based tendencies of a modal verb of necessity.

Acknowledgments

References

Arkisyn. 2018. *A morphosyntactically coded database of conversational Finnish*. Database 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.

Bates, Douglas, Martin Maechler, Ben Bolker & Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1). 1–48. https://doi.org/10.18637/jss.v067.i01.

Bybee, Joan & Sandra Thompson. 1997. *Three frequency effects in syntax*. Berkeley Linguistics Society 23. 65–85. https://doi.org/10.3765/bls.v23i1.1293.

Clark, H. H. 1973. The language-as-fixed-effect fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behavior* 12(4). 335–359. https://doi.org/10.1016/S0022-5371(73)80014-3.

Couper-Kuhlen, Elizabeth & Marja Etelämäki. 2015. Nominated actions and their targeted agents in Finnish. *Journal of Pragmatics* 78, 7–24. https://doi.org/10.1016/j.pragma.2014.12.010.

Dryer, Matthew S. 2013. Expression of pronominal subjects. In Matthew S. Dryer & Martin Haspelmath (eds.), *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available at http://wals.info/chapter/101, accessed 2022-01-16)

Goodwin, Charles. 1984. Notes on story structure and the organization of participation. In J. M. Atkinson & John Heritage (eds.), *Structures of Social Action: Studies in Conversation Analysis*, 225–246. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511665868.016.

Goodwin, Charles & Marjorie H. Goodwin. 1990. Context, activity and participation. In Peter Auer & Aldo di Luzo (eds.), *The contextualization of language*, 77–99. Amsterdam: Benjamins. https://doi.org/10.1075/pbns.22.07goo.

Gries, Stefan Th. 2021. *Statistics for Linguistics with R: A Practical Introduction*. 3rd. edition. Berlin/Boston: De Gruyter Mouton.

Gundel, Jeanette K., Nancy Hedberg & Ron Zacharski. 1993. Cognitive status and the form of referring expressions in discourse. *Language* 69(2). 274–307. https://doi.org/10.2307/416535.

Hakulinen, Auli. 1998. The use of Finnish *nyt* as a discourse particle. In Andreas Jucker & Yael Zif (eds.), *Discourse particles. Descriptions and Theory*, 83–96. Amsterdam: John Benjamins. https://doi.org/10.1075/pbns.57.06hak.

Hakulinen, Auli. 2001. On some uses of the discourse particle *kyllä* in Finnish conversations. In Selting, Margret, Elizabeth Couper-Kuhlen & Sandra A. Thompson (eds.), *Studies in Interactional Linguistics*, 10, 171–198. Philadelphia: John Benjamins.

Hakulinen, Auli & Lauri Karttunen. 1973. Missing persons: On generic sentences in Finnish. In Claudia W. Corum, T. Cedric Smith-Stark & Ann Weiser (eds.), *Papers from the Ninth Regional Meeting of the Chicago Linguistic Society*, 157–171. Chicago, IL: Chicago Linguistic Society.

Hakulinen, Auli, Maria Vilkuna, Riitta Korhonen, Vesa Koivisto, Tarja Riitta Heinonen & Irja Alho. 2004. *Iso suomen kielioppi*. Helsinki: Finnish Literature Society. http://scripta.kotus.fi/visk, URN: ISBN:978-952-5446-35-7 (accessed 7 January 2021).

Heinonen, Tarja Riitta. 1995. Null subjects in Finnish. From either-or to more-or less. In Tapio Hokkanen, Marja Leinonen & Susanna Shore (eds.), *SKY Journal of Linguistics* 1995, 47–78. Helsinki: Suomen kielitieteellinen yhdistys.

Helasvuo, Marja-Liisa. 2014. Jotta suomalaiset voisivat puhua enemmän. Puhetilanteen osallistuja tekstiviestikeskustelussa. In Marja-Liisa Helasvuo, Marjut Johansson & Sanna-Kaisa Tanskanen (eds.), *Kieli verkossa: Näkökulmia digitaaliseen vuorovaikutukseen*, 29–49. Helsinki: Finnish Literature Society.

Helasvuo, Marja-Liisa & Lea Laitinen. 2006. Person in Finnish. Paradigmatic and syntagmatic relations in interaction. In Marja-Liisa Helasvuo & Lyle Campbell (eds.), *Grammar from the human perspective. Case, space and person in Finnish. Current issues in linguistic theory*, 277, 173–207. Amsterdam: John Benjamins. https://doi.org/10.1075/cilt.277.14hel.

Helasvuo, Marja-Liisa & Maria Vilkuna. 2008. Impersonal is personal: Finnish perspectives. *Transactions of the Philological Society* 106(2). 216–245. https://doi.org/10.1111/j.1467-968X.2008.00208.x.

Helasvuo, Marja-Liisa & Aki-Juhani Kyröläinen. 2016. Choosing between zero and pronominal subject: Modeling subject expression in the 1st person singular in Finnish conversation. *Corpus Linguistics and Linguistic Theory* 12(2). 263–299. https://doi.org/10.1515/cllt-2015-0066.

Helasvuo, Marja-Liisa. 2019. Free NPs as units in Finnish. *Studies in Language* 43(2). 301–328. https://doi.org/10.1075/sl.16064.hel.

Hopper, Paul. 1987. *Emergent grammar.* Berkeley Linguistics Society 13. 139–157. https://doi.org/10.3765/bls.v13i0.1834.

Kehayov, Petar & Reeli Torn-Leesik. 2009. Modal verbs in Balto-Finnic. In Björn Hansen & Ferdinand de Haan (eds.), *Modals in the languages of Europe: A reference work*, 363–402. Berlin/Boston: Mouton de Gruyter.

Jaakola, Minna. 2004. *Suomen genetiivi*. SKST 995. Helsinki: SKS.

Jokela, Hanna. 2012. *Nollapersoonalause suomessa ja virossa. Tutkimus kirjoitetun kielen aineistosta*. Scripta Lingua Fennica Edita 334. Turku: University of Turku.

Jokela, Hanna & Helen Plado. 2015. Subjects under generic conditions. Implied subjects in Finnish and Estonian if-clauses. In Marja-Liisa Helasvuo & Tuomas Huumo (eds.), *Subjects in constructions – canonical and non-canonical: canonical and non-canonical*, 43–72. Amsterdam: John Benjamins Publishing Company. https://doi.org/10.1075/cal.16.04jok.

Kangasniemi, Heikki. 1992. *Modal expressions in Finnish*. Studia Fennica Linguistica 2. Helsinki: Finnish Literature Society.

Koivisto, Aino, Ritva Laury & Eeva-Leena Seppänen. 2011. Syntactic and actional characteristics of Finnish *että*-clauses. In R. Laury & R. Suzuki (eds.), *Subordination in conversation: A cross-linguistic perspective*, 69–102. Amsterdam: John Benjamins. https://doi.org/10.1075/slsi.24.05koi.

Laitinen, Lea. 1995. Nollapersoona. *Virittäjä* 99. 337–358.

Laitinen, Lea. 2006. Zero person in Finnish: A grammatical resource for construing human reference. In Marja-Liisa Helasvuo & Lyle Campbell (eds.), *Grammar from the human perspective: case, space, and person in Finnish*, 209–233. Amsterdam: Benjamins. https://doi.org/10.1075/cilt.277.15lai.

Lindström Jan, Camilla Lindholm & Ritva Laury. 2016. The interactional emergence of conditional clauses as directives: Constructions, trajectories and sequences of actions. *Language Sciences* 58, 8–21. https://doi.org/10.1016/j.langsci.2016.02.008.

Lindström, Liina, Mervi Kalmus, Anneliis Klaus, Liisi Bakhoff & Karl Pajusalu. 2009. Ainsuse 1. isikule viitamine eesti murretes. *Emakeele Seltsi aastaraamat* 54, 159–185.

Nielsen, Søren Beck, Christina Fosgerau, Jensen Fogtmann & Juel Torben. 2009. From community to conversation – and back: Exploring the interpersonal potentials of two generic pronouns in Danish. *Acta Linguistica Hafniensia* 41(1), 116–142. https://doi.org/10.1080/03740460903364151.

Pajusalu, Renate. 2015. Hidden subjects in conversation. Estonian personless verb forms as referential devices. In Marja-Liisa Helasvuo & Tuomas Huumo (eds.), *Subjects in constructions – canonical and non-canonical: canonical and non-canonical*, 43–72. Amsterdam: John Benjamins Publishing Company. https://doi.org/10.1075/cal.16.03paj.

Ragnarsdottir, Hrafnhildur & Sven Strömqvist. 2005. The development of generic maður/man for the construction of discourse stance in Icelandic and Swedish. *Journal of Pragmatics* 37, 143–155. https://doi.org/10.1016/S0378-2166(04)00194-8.

R Core Team. 2020. *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Sacks, H., Emanuel A. Schegloff & Gail Jefferson. 1974. A simplest systematics for the organization of turn-taking for conversation. *Language* 50(4), 696–735. https://doi.org/10.1353/lan.1974.0010.

Siewierska, Anna. 2008. Ways of impersonalizing: Pronominal vs verbal strategies. In María de los Ángeles Gómez González, J. Lachlan Mackenzie & Elsa M. González Álvarez (eds.), *Current trends in contrastive linguistics: Functional and cognitive perspectives (Studies in Functional and Structural Linguistics 60)*, 3–26. Amsterdam: John Benjamins. https://doi.org/10.1075/sfsl.60.03sie.

Sorjonen, Marja-Leena. 2001. *Responding in conversation: A study of response particles in Finnish.* Philadelphia: John Benjamins. https://doi.org/10.1075/pbns.70.

Travis, Catherine. 2005. The yo-yo effect: Priming in subject expression in Colombian Spanish. R. Gess & E. Rubin (eds.), *Theoretical and experimental approaches to Romance linguistics: Selected papers from the 34th Linguistic Symposium in Romance Languages* 2004, 329–349. Amsterdam: John Benjamins. https://doi.org/10.1075/cilt.272.20tra.

Uusitupa, Milla. 2021. When Finnish and Karelian person systems come together: Comparing open person constructions in Border Karelian dialects. *Eesti ja soome-ugri keeleteaduse ajakiri. Journal of Estonian and Finno-Ugric Linguistics* 12(1). 397–431. https://doi.org/10.12697/jeful.2021.12.1.11.

Varjo, Mikael. 2019. It takes all kinds to make a zero: Employing multiple correspondence analysis to categorize an open personal construction in Finnish. *Corpus Linguistics Research* 5, 55–87. https://doi.org/10.18659/CLR.2019.5.03.

Varjo, Mikael & Karita Suomalainen. 2018. From zero to 'you' and back: A mixed methods study comparing the use of two open personal constructions in Finnish. *Nordic Journal of Linguistics* 41(3), 333–366. https://doi.org/10.1017/S0332586518000215.

Vepsäläinen, Heidi. 2019. *Suomen no-partikkeli ja kysymyksiin vastaaminen keskustelussa.* Helsinki: Helsingin yliopisto. https://doi.org/10.23982/vir.82612.

Vilkuna, Maria. 1989. *Free word order in Finnish. Its syntax and discourse functions*. Helsinki: SKS.

Väänänen, Milja. 2016. *Subjektin ilmaiseminen yksikön ensimmäisessä persoonassa. Tutkimus suomen vanhoista murteista*. Turun yliopiston julkaisuja C430. Turku: Turun yliopisto. http://urn.fi/URN:ISBN:978-951-29-6664-6.

Winter, Bodo. 2013. *Linear models and linear mixed effects models in R with linguistic applications.*

Winter, Bodo & Martine Grice. 2021. Independence and generalizability in linguistics. *Linguistics* 59(5). 1251–1277. https://doi.org/10.1515/ling-2019-0049.

Zinken, Jörg & Eva Ogiermann. 2011. How to propose an action as objectively necessary: The case of Polish *trzeba x* ("one needs to x"). *Research on Language and Social Interaction* 44(3). 263–287. https://doi.org/10.1080/08351813.2011.591900.

Zuur, Alain F., Elena N. Ieno & Chris S. Elphick. 2010. A protocol for data exploration to avoid common statistical problems. *Methods in Ecology and Evolution* 1(1), 3–14. https://doi.org/10.1111/j.2041-210X.2009.00001.x.

Kokkuvõte. Mikael Varjo: Enam kui null? Referentsiaalselt erilised ja avatud netsessiivkonstruktsioonid soome argivestlustes. Käesolevas artiklis uuritakse grammatilise subjekti vaheldumist. Võrdlen omavahel soomekeelses argivestluses esinevaid *pitää*-verbiga moodustatud modaalseid netsessiivkonstruktsioone, mis esinevad kas subjektita või sisaldavad kõneaktis osalevat subjekti positsioonis pronoomenit (pronoomensubjekti). Soome keeles tähistatakse grammatilist subjekti tüüpiliselt nii pronoomeni kui ka pöördelõpuga, kuigi põhimõtteliselt võib subjekti funktsioonis pronoomeni ära jätta. Netsessiivverbid on siiski vaegpöördelised, esinedes ainult ainsuse 3. pöörde vormis, mis tähendab seda, et verb ei ühildu subjektiga. Netsessiivkonstruktsiooni puhul on tavaline pronoomensubjekti täielik ärajätt, mille tulem on nn üldisikuline konstruktsioon (*nollapersoonarakenne*) – soome keelele omane referentsiaalselt avatud isikut väljendav konstruktsioon. Üldistatud lineaarse segamudeli (GLMM) tulemused näitavad, et subjektivaheldust mõjutavad mitmed semantilis-grammatilised ning ka diskursiivsed tegurid. Tulemused näitavad, et üldisikulistel netsessiivkonstruktsioonidel on selgelt eristatavad sotsiaalset interaktsiooni liigendavad funktsioonid, näiteks (ühine) planeerimine või soovide väljendamine ja kõnesituatsiooni osalejarollide muutmine. Grammatilist subjekti sisaldavad netsessiivkonstruktsioonid aga rõhutavad pigem teo või tegevuse vajalikkust või kohustuslikkust kui selle otsest algatamist.

Märksõnad: korpuslingvistika, soome keel, suhtluslingvistika, netsessiivkonstruktsioon, avatud viitesuhe, süntaks

Appendix

Transcription symbols

.	falling intonation
,	level intonation
?	rising intonation
↑	step up in pitch
↓	step down in pitch
sp<u>ea</u>k	emphasis
>speak<	faster pace than in the surrounding talk
<speak>	slower pace than in the surrounding talk
°speak°	quiet talk
sp-	word cut off
spea:k	lengthening of a sound
#speak#	creaky voice
£speak£	smiley voice
.h	audible inhalation
h	audible exhalation
.speak	word spoken during inhalation
[	beginning of overlap
]	end of overlap
=	latching of units
(.)	micropause (less than 0.2 seconds)
(0.6)	pause length in tenth of a second
(speak)	item in doubt
(-)	item not heard
boldface	focused item in the transcript