SEARCHING FOR MOTIVATIONS FOR GRAMMATICAL PATTERNINGS

Marja-Liisa Helasvuo

Abstract

In this article I analyze subject expression in conversational Finnish, identifying the home environments for zero and pronominal subjects in the 1st and 2nd person singular. Based on a syntactically coded database, I show that there is a clear preference, in both 1st and 2nd person, for pronominal subjects over zeros; in other words, double-marking is preferred over single-marking. This clearly contravenes the general preference for minimization or economy in person reference in conversation, as suggested by Sacks and Schegloff (1979) and Levinson (2007; see also Hacohen and Schegloff 2006). The home environments for zero and pronominal subjects are analyzed in terms of the micro-level social actions performed by participants, in order to find motivations for the choice of the form of subject. The analysis of the Finnish data shows that the choice between zero vs. pronominal subject is sensitive to features in the sequential context. It affects turn projection.

The article shows that a systematic analysis of the data can provide important insights regarding global patterns. The deeper motivations that lie behind these patternings, however, cannot be understood without close microanalysis of the local contexts of subject expression.

Keywords: Subject expression; Subject omission; Person marking; Economy; Projection.

1. Introduction

According to Sacks and Schegloff (1979), there is a general preference in conversation for minimization in reference to person: Reference to person is "preferredly done with a single reference form". Moreover, there is a general preference for recipient design: Speakers should use reference forms that allow the recipients to recognize who is being referred to (Sacks & Schegloff 1979: 16). Levinson (2007) discusses "optimizing" in reference to person based on data from Yélî Dnye, Rossel Island. He notes that there are several – sometimes conflicting – principles at work, of which three are the most important: Economy (cf. Sacks and Schegloff's minimization), recognition (cf. recipient design), and circumspection. According to the recognition principle, speakers should restrict the set of referents so as to achieve recognition. The economy principle says that speakers should minimize expressive means by using a single referring expression so as to avoid over-restricting the set of referents explicitly. Finally, according to the principle

of circumspection, speakers should not over-reduce the set of referents.[1] In broad terms, the principles of both economy and circumspection aim at reference that is sufficient for the needs of the participants. In this article, I discuss these principles in the light of data from conversational Finnish. I focus on subject expressions.

In Finnish, the predicate verb agrees with the subject in number (singular vs. plural) and person (1st, 2nd and 3rd; see e.g. Sulkala & Karjalainen 1992). In standard Finnish, non-3rd-person subjects are generally not expressed (i.e. zero anaphora is the norm for 1st and 2nd person), while the verb is marked for person and number of the subject.[2] In other words, standard Finnish follows the principle of minimization/economy.[3] When the subject is overtly expressed, this serves some special discourse function, such as, for example, contrast. The system of standard Finnish is thus similar to that of conversational Hebrew (see Hacohen and Schegloff 2006; Ariel 1990: 48–49): While subject pronouns are typically elided and person reference is conveyed through agreement marking on the verb,[4] double-marking (i.e. subject pronoun + verbal person marking) is also possible; it is used to perform micro-level actions, such as marking the current utterance as dispreferred, as in Hebrew (cf. Hacohen & Schegloff 2006), or as contrastive, as in Finnish.

In casual conversational Finnish, however, double-marking is the norm in the expression of 1st and 2nd-person reference (see Helasvuo & Laitinen 2006). In other words, verbal person markers are obligatory and subject pronouns are usually overtly expressed (see ex. (1)). (See section 2 for a closer description of the data.)

(1) (D113a)

1 Sari: *kuitenki sillo joskus   mä   muista-n*
          anyway then sometime 1SG remember-1SG
          'Anyways, I remember one time'

2   *mä   meni-n   linka-lla (.) Juha-n   luo   sitte*
          1SG go-1SG bus-ADE   MaleName-GEN towards then
          'I went to Juha's place with the bus'

In (1), subjects are expressed by the 1st-person singular pronoun *mä* (lines 1 and 2). The finite verbs *muistan* '(I) remember' and *menin* '(you) went' index the respective subject with the person suffix (*-n*).

---

[1] There is a rich literature on person reference and its relation to deixis and accessibility (see e.g. Ariel 1990; Levinson 2004). Kibrik (2009) is a comprehensive overview of reference in discourse from a cross-linguistic perspective. This paper has a more limited scope as it focuses on person reference in conversation.

[2] In the 3rd person, pronominal subjects are generally expressed in both standard and colloquial Finnish (cf. Sulkala & Karjalainen 1992), but can be left out in certain anaphoric contexts (Hakulinen & Laitinen 2008).

[3] By standard Finnish, I refer to formal written and spoken Finnish. There is a norm concerning the use of pronominal subjects in the 1st and 2nd person. It developed gradually in the course of several centuries (starting in the 17th century). It was debated in prescriptive writings especially in the 19th century. Among the arguments presented to support the avoidance of pronominal subjects, economy and avoidance of redundant markings were mentioned. (Strellman 2005.)

[4] In Hebrew, this applies to person reference in past and future tenses (see Hacohen and Schegloff 2006), in Finnish to all tenses.

It is interesting to note that the preference for minimization in referring to persons, as described by Sacks and Schegloff (1979), and Levinson (2007), does not hold in conversational Finnish, where there rather appears to be a preference for double-marking in the 1st and 2nd person. Single-marking (minimization) is also possible in colloquial Finnish, especially in certain conversational contexts (such as in the answer part of a question–answer adjacency pair; for preliminary studies on the home environments for subject omission in everyday conversation see Duvallon 2006, and in institutional talk see Lappalainen 2006). In sum, both double-marking and single-marking (verbal person marking only) are grammatically possible in conversational Finnish, but the expression vs. omission of subject is contextually constrained.

For English, where overt expression of the pronominal subject (with the exception of same-subject coordination and imperatives) is the norm, it has been suggested that when main-clause subjects in English are not expressed (i.e. in cases of zero anaphora), speakers are using zero anaphora as a resource to perform specific micro-level social actions, such as marking the current utterance as a "resaying" or "secondary action" (Oh 2005: 296) or as "resuming prior turn-constructional unit after parenthetical insert" (Oh 2006: 822–830). With regard to Finnish, Lappalainen (2006) has suggested that in certain institutional settings subject omission may be related to epistemic stance.

In this article, I use a syntactically coded database of conversational Finnish to examine the contexts for both subject expression and subject omission, in order to explicate the home-environments for zero and pronominal subjects. I analyze these contexts in terms of the micro-level social actions performed by participants, searching for motivations for the choice of the referential form of the subject.[5] I show that in both the 1st and the 2nd person there is a clear preference for pronominal subjects over zeros; in other words, double-marking is preferred over single-marking. This preference, however, is considerably stronger in the 1st person than in the second. I report the results of the analysis for several contextual features, and discuss motivations for the patternings observed.

After describing my data and the coding scheme for the database, I proceed to a discussion of the home environments for zero subjects. This is followed by an examination of typical contexts for pronominal subjects. Finally, I discuss questions as a special environment for subject expression.

2. The data

The data for this study come from eleven face-to-face conversations in Finnish between family and friends. Altogether the data comprise more than seven hours of recordings. The data come from the Spoken Language Archives at the University of Turku and the Conversation Analysis Archives at the University of Helsinki.

In the initial analysis, the data were segmented into syntactic units: Clauses, free/unattached NPs or particles forming utterances of their own (see Helasvuo 2001: 21–33, 105–113). In Finnish, only nominative subjects can trigger person agreement in the verb. This means that only clauses with nominative subject have the possibility of

---

[5] For the purposes of this article, the term *subject expression* refers to 1st and 2nd person pronominal subjects and *subject omission* to zero subjects. Note that in the case of zero subjects, there is person marking on the verb, but the subject is omitted.

double-marking, i.e. person marking through both the pronominal subject and the verbal person marking. From the data set, clauses which in principle could have a nominative subject were extracted for further study. These include transitive, intransitive and predicate nominal clauses. The total number of clauses extracted for this study is 7862. The quantitative findings presented in this paper are based on these data, with a special focus on 1st and 2nd person singular clauses (N = 2231). The qualitative analysis is based on the recordings of the actual conversations and the transcripts.

The data were further analyzed and coded for several features. Some codings concerned the formal properties of the subject, while others dealt with reference and referential distance. Some properties of the finite verb were also coded, as well as some clausal features. Of the formal properties of the subject, its structure was coded: i.e. whether the subject was in pronominal or zero form. If an overt subject was present, its distance from the predicate verb was coded (adjacent to the verb vs. separated by intervening elements). With regard to referential features, the distance of the last mention of the subject referent – if there was one – was coded. Predicate verbs were analyzed in terms of tense and the syntactico-semantic type of the verb. The polarity (affirmative vs. negative) of the clause was coded, as was the presence and form of a possible object NP.

In this article I focus on subject expression and omission in the first and second person singular. Table 1 summarizes the distribution of pronominal vs. zero subjects across 1st and 2nd person singular verb forms in the data.

Subject expression	Verb form				Total	
	SG1		SG2[*]			
	N	%	N	%	N	%
Zero	203	11	101	24	304	14
Pronoun	1590	89	327	76	1927	86
Total	1793	100	428	100	2231	100

Table 1. *Distribution of pronominal vs. zero subjects with 1st and 2nd person singular verb forms.* (* imperatives excluded.)

As we see from Table 1, a pronominal subject is favored over zero in both the 1st and the 2nd person singular; 86% of the clauses with either 1st or 2nd person singular subjects have a pronominal subject. We also see, however, that this preference is less strong in the 2nd person, where 24% of clauses have zero subjects. I return to this difference between the 1st and 2nd person later in the article. The findings shown in Table 1 receive additional support when compared to Lappalainen (2004), who has studied the expression of 1st and 2nd person subjects in Finnish. Her data, representing ordinary conversation, showed a separate subject pronoun in addition to the verbal person suffix in a majority of 1st person singular clauses (90% or 422/471); in 2nd person singular clauses the proportion of pronominal subjects was somewhat smaller (79 % or 151/190; see Lappalainen 2004: 81). In other words, Lappalainen's findings are very close to those shown in Table 1. It should be noted that Table 1 does not include imperative forms. Second person imperatives are overwhelmingly used without

a subject.[6] Moreover, 2nd person imperative verb forms do not have person suffixes. Both facts make them quite exceptional in terms of subject expression, and they are thus excluded from the discussion for the remainder of this article.

3. Home environments for subject omission

As noted above, there is a preference in conversational Finnish for double-marking in subject expression in the 1st and 2nd person singular. Single-marking is also possible, but its use is more limited. Several studies have suggested that there are certain conversational contexts which favor single-marking. Duvallon & Chalvin (2004) and Duvallon (2006) have studied subject omission in 1st and 2nd person singular forms. Both Duvallon & Chalvin and Lappalainen (2004) base their studies on various varieties of spoken interaction, ranging from ordinary conversation among friends and family to various institutional settings. Lappalainen (2006a and 2006b) focuses on certain institutional contexts. In this section, I discuss typical contexts for subject omission, trying to identify the home environments for subject omission in everyday conversation.

Duvallon (2006; see also Duvallon & Chalvin 2004) identifies complex structural units as one of the home contexts for subject omission. Complex structural units include adjacency pairs containing a repeat of the finite verb and cases where an identical subject pronoun is omitted in the latter part of the complex unit (e.g. in same-subject coordination). For example, in question–answer pairs it is possible in Finnish to respond to a polar question with the finite verb alone (2); see Sorjonen (2001a).

(2) (SG151)

1 Susa: *nii oo-t   sie   jo   tä-nä   aamu-na   ol-lu,*
           so be-2SG 2SG already this-ESS morning-ESS be-PCP

2   *jo   sali-lla*
           already gym-ADE
           'So have you already been at the gym this morning?'

3 Anu: *oo-n.*
           be-1SG
           'I have.'

In (2) Susa uses a 2nd person singular verb form and a pronominal subject, while Anu produces just the finite part of the verb in her response, with the 1st person singular verbal inflection. Had Anu used the pronominal subject in her response, the response would not have served as a neutral affirmative answer to the question, but rather

---

[6] In the data for this study, there were 164 second person imperative forms, and only three of them appeared with the second person pronoun in the nominative case, i.e. with a possible subject candidate. In the research literature, there is controversy over the status of the possible pronoun: Some researchers argue for analyzing it as a subject, some say it is an apposition (see e.g. Hakulinen et al. 2004: 1565). Even if the pronoun was analyzed as a subject, it is obvious that it is very different from other subjects and should be treated separately.

contrasting the referent of the 1st person singular pronoun with something else (e.g. 'I have but somebody else hasn't').

In coordinated clauses with same subject, the subject pronoun is omitted in the latter clause; see (3). Similarly, in self-repair targeted at the verb form the subject pronoun is not repeated, as shown in (4).

(3) (D113b)

1 Ben: *kumma-ssa tapaukse-s   sä   kuol-isi-t*
            which-INE   case-INE   2SG   die-CON-2SG
            'In which case would you die,'

2   *(.) sä   jä-isi-t   tähän vai*
              2SG   stay-CON-2SG   here   or
            '(if) you stayed here or'

3   *läht-isi-t   juokse-ma-han*
            leave-CON-2SG   run-INF-ILL
            '(if) (you) started running.'

(4) (D113b)

1 Ben: *kun   mä (.) men-i-n   [lähr-i-n*
            when 1SG go-PST-1SG   leave-PST-1SG
            'when I went left'

2 Mikko: [*luje-mpa-a*
                                          fast-CMP-PAR
                                          'faster'

3   *mitä   vitosella*
            than   five-ADE
            'than with gear 5.'

4 Ben: *lähr-i-n   aja-ma-han*
            leave-PST-1SG   drive-INF-ILL
            'started driving'

In (3), Ben asks a wh-question in line 1, and then gives two alternative answers (lines 2 and 3). The alternatives are presented in a coordinated clausal compound. In the first part of the compound (line 2), Ben uses both the subject pronoun and the verbal person suffix (*sä jäisit* '(if) you stayed'), but in the latter part of the compound, the subject pronoun is not repeated (*lähtisit juoksemahan* '(if) (you) started running'). In (4), in line 1 Ben makes a self-repair, replacing the verb *menin* 'I went' with *lährin* 'I left'. In making the repair, the subject pronoun *mä* 'I' is not repeated; rather, just the finite verb form is produced. The repair is produced in overlap with Mikko (line 2), and the verb form is repeated in line 4, also without the subject pronoun, when the overlapping turn is finished.

In addition to complex structural units, Duvallon (2006) identifies modality directed to the addressee as one of the home contexts for subject omission. In certain contexts, statements can be used as directives, as in (5). In directive usage the subject is typically omitted. Likewise optatives (i.e. utterances expressing a wish) concerning the co-participant typically occur without the subject pronoun; see (6).

(5) (C158a)

J: *Mut ne   ryäppää-t sitte*
            but 3PL blanch-2SG then
            'but those you should blanch'

            neh (.) karvalaukut (.)
            3PL   milk cap-PL
            'those milk caps (mushrooms)'

(6) (D113b)

Ben: *voi voi   sa-isi-t   myy-rä mu-lle*
            PTC PTC shall-CON-2SG sell-INF 1SG-ALL
            'Oh you should sell me'

            se-n   kakskymppise-n
            DEM-ACC   twenty-ACC
            'that 20 (hard drive).'

In (5), J is giving advice to his daughter on how to prepare mushrooms. The directive is formally a declarative. There is a person suffix *–t* on the verb *ryäppäät* '(you) blanch' but no subject pronoun (the preverbal 3rd person plural pronoun functions as an object). The directive meaning of the utterance is contextually induced and reinforced by the lack of subject pronoun. In ex.
(6), Ben expresses a wish concerning the future activities of the co-participant. He does so by means of the 2nd person conditional form of the main verb *saisit* '(you) should' without a pronominal subject. The optative meaning is conveyed through the lack of a subject pronoun and the conditional verb form.

As mentioned earlier, Levinson (2007) proposes several conversational principles guiding the choice of reference forms in interaction, the most important of which are the principles of economy, recognition, and circumspection. In general, subject omission seems to work in accordance with the principle of economy or minimization, but they do not seem like plausible motivating factors, since cases with single-marking are clearly in the minority. A closer analysis of the home environments for subject omission reveals that the environments respond to the principles differently. In environments that form complex structural units – in the response part of a yes/no question–answer sequence, in self-repairs targeted at the verb, and in same-subject coordination – subject omission is in accordance with the economy principle. Furthermore, it follows the principle of recipient design (cf. Schegloff 1979): The larger sequence within which the construction with the omitted subject is embedded functions to achieve early recognition and to allow the co-participant to project the trajectory of the turn and the larger sequence. The term projection refers to the various interactional practices that

foreshadow the future course of the turn and thus enable the recipients to anticipate turn transition (see e.g. Clayman 2012). Example (2) shows an overt pronominal subject in the question part of the question–answer adjacency pair, and (3) has a pronominal subject in the first part of the coordination (line 2). In the self-repair in (4), the speaker backs up to the finite verb but does not repeat the pronominal subject. In other words, in all of these environments there is an overt pronominal reference to the referent of the subject in the context immediately prior to the omitted subject. This allows for early projection of the future course of the interactional sequence.

In directives and optatives (examples (5) and (6)) early projection is achieved not on the basis of the linguistic context but on that of the immediate speech situation: It is to the addressee that directives such as (5) can be successfully directed. We may further note that modality directed to the addressee is a context where second person forms are typically used; in the case of directives, first person forms are typically not used at all (see e.g. König & Siemund 2007: 21). This may explain the finding discussed in connection with Table 1: We noted that although both 1st and 2nd person favor pronominal subjects over zeros, this preference is stronger in the 1st person. For the 2nd person, there are more environments that favor zero subjects than for the 1st person.

In sum, despite the overall preference for double marking in person reference, it is possible to use single-marking for specific discourse purposes. Moreover, single-marking is the preferred alternative in many contexts, such as in the answer part of a question–answer adjacency pair (cf. (3)) and in statements used as directives (cf. (5)) or optatives (cf. (6)). These contexts are characteristic for Finnish subject omission. In addition, there are certain contexts in which even languages that "normally if not obligatorily" have pronouns in subject position, such as English (see Dryer 2011: Chapter 101), allow for subject omission. Same-subject coordination (cf. (3)) is one such context. Both economy and minimization in reference to person have been proposed as motivations for single-marking (cf. Schegloff 1979; Hacohen & Schegloff 2006; Levinson 2007). The present data indicate that instead of minimization or economy, recipient design is a far more important factor guiding the choice of referential form: Subject omission is preferred in contexts where the larger sequential context allows for early recognition and projection of the future trajectory of the turn. As we will see in the next section, early projection is important in the environments for subject expression as well.

4. Home environments for subject expression

While there have been some studies exploring the typical contexts for subject omission in Finnish (see previous section), the home environments for pronominal subjects have received much less attention, despite the fact that pronominal subjects are by far the more frequent alternative (cf. Table 1 above). The current database shows that pronominal subjects favor cognitive verbs in the 1st and 2nd person. I argue that verbs of cognition form emergent discourse patterns together with 1st and 2nd person pronominal subjects, and I discuss typical contexts in which these patterns are used and the conversational actions they are used to perform. The database further indicates that pronominal subjects are likely to occur in contexts where the referent of the pronoun has

not been mentioned in the immediately prior talk. I will now disucss these contexts in detail.

Discourse studies have indicated that there is a general, cross-linguistic tendency for 1st and 2nd person subjects to co-occur with verbs of cognition. In Scheibman's (2002: 63) data representing American English conversation, 1st person subjects appeared most often with verbs of cognition (32% of 1st person singular subjects; see also Tao 2001; Kärkkäinen 2003 and 2007 for American English; Kaltenböck 2007 for British English; Tao 1996: 152 for Mandarin; Weber and Bentivoglio 1991; and Torres Cacoullos & Travis 2011 for Spanish). There are also several studies indicating that in English, 1st and 2nd person subjects form such regular and frequent combinations with certain cognition verbs that they crystallize into fixed units, such as *I think*, *you know*, *I mean* etc. (see e.g. Kärkkäinen 2003; Östman 1981). Thompson and Mulac (1991) call these epistemic parentheticals, and Kärkkäinen (2003 and 2007) argues that the epistemic phrase together with the associated utterance performs a certain stance-taking action in interaction. In a study on Estonian interaction, Keevallik (2003) found a close correlation between 1st person subjects and verbs of cognition (such as *teadma* 'know' and *arvata* 'think'). She further suggests that in Estonian the phrase *mai tea ~ ma ei tea* 'I don't know' has features of "a disaligning or disjunctive particle or epistemic adverb" (Keevallik 2003: 98). Helasvuo (2001a) has suggested the possible occurrence of similar crystallization processes in Finnish too.

The current database shows that the preference for verbs of cognition is particularly clear for 1st person singular pronominal subjects. Table 2 focuses on 1st person singular subjects comparing their distribution across different types of verbs. The analysis of verb types is based on their semantics and argument structure (see Pajunen 2001; Dixon 2005: Part B).

Verb type	Subject expression				Total
	zero		pronoun		
	N	%	N	%	N
action, activity and change	23	12	164	88	187
cognition	48	8	517	92	565
motion	63	15	355	85	418
perception	21	15	120	85	141
physiological state or process	7	11	56	89	63
secondary	0	0	8	100	8
speech act	21	10	187	90	208
state	20	10	183	90	203
Total	203	11	1590	89	1793

Table 2. Type of subject (zero vs. pronoun) across verb type in the 1st person singular.

Table 2 shows that pronominal subjects are preferred over zeros in all verb types[7]. This tendency is particularly clear in verbs of cognition. Statistical analysis of the data presented in Table 2 shows that there is a statistically significant positive association

---

[7] Verbs describing emotional states and processes are also included under "verbs of cognition". Modal verbs and the negation verb *ei* were coded as "secondary verbs" if used alone, without the main verb, either the infinitive or the connegative form of the verb (see Dixon 2005: 96-101).

between verbs of cognition and pronominal subjects, and a statistically significant positive association between motion verbs and zero subjects.[8]

We will now look at the most frequent verbs of cognition and their distribution in connection with different verbal person forms. The most frequent verbs of cognition in the data are *tietää* 'know', *ajatella* 'think', and *muistaa* 'remember'. All of these verbs strongly prefer singular 1st person forms over other forms of verbal person inflection, as can be seen in Table 3.

Verb	Verbal person form											
	SG1		SG2		SG3		PL2		Pass.		Total	
	N	%	N	%	N	%	N	%	N	%	N	%
ajatella 'think'	90	68	18	14	14	11	2	1	8	6	132	100
haluta 'want'	41	50	13	16	23	28	2	2	2	4	81	100
luulla 'think, believe'	32	65	5	10	10	20	0	0	2	4	49	100
muistaa 'remember'	91	82	14	13	5	4	1	1	0	0	111	100
tietää 'know'	177	64	35	13	52	19	8	3	6	2	278	100
Total	431	66	85	13	104	16	13	2	18	3	652	100

Table 3. Five most frequent verbs of cognition and their distribution across different person forms.

The special nature of the relationship between verbs of cognition and first and second person verb forms becomes even clearer if we compare the figures in Table 3 to the overall distribution of verbs across different person forms: In the data as a whole, singular forms are more common than plural forms (88% of finite verbs are in the singular), but within the singular forms, it is the 3rd person singular form which is by far the most frequent, accounting for 65% of all singular verb forms. This can be compared to the five most frequent verbs of cognition shown in Table 3: Only 16% (104/652) of these verbs of cognition are in the 3rd person singular and 66% (431/652) are in the 1st person singular.

A closer microanalysis of the contexts where the most frequent verbs of cognition (see Table 3) occur reveals that each of these verbs has its own profile. I illustrate these profiles with the help of examples of *tietää* 'know', *ajatella* 'think', *muistaa* 'remember' and *luulla* 'think, believe'. I have searched for interesting patterns with the help of the database and then made collections of examples which have been further

---

[8] A Pearson Chi-Square test was performed on the data presented in Table 2, $X^2(7, N = 1793) = 14,09$, p < ,05. The strength of the association in individual cells was estimated with standardized Pearson residuals. Any residual higher than 2 or lower than -2 shows that the association is statistically significant. The relevant residuals are as follows: Pronominal subjects occurring with verbs of cognition, adjusted residual 2,6; pronominal subjects occurring with motion verbs, adjusted residual -2,8; zero subjects occurring with motion verbs, adjusted residual 2,8. The other cells did not show associations that would be statistically significant.

analyzed in their natural contexts. Let us first consider the typical contexts for *tietää* 'know', as shown in (7) below.

(7) (C133)

1 EK:	*kyl tei-l o k- Malle kauhean kiva*
	PTC 2PL-ADE be-3SG Name awfully nice
	'well it sure is neat Malle'

2	*se teiä-n kesä-paikka on.*
	DEM 2PL-GEN summer-place be+3SG
	'that summer place you have'

| 3 | (.) |

4 MS:	*nii >e-m mä tiä<*
	PTC NEG-1SG 1SG know-CONNEG
	'well I don't know'

5	*kyl se, (.) kyl se >semmone ku*
	PTC DEM PTC DEM DEM-ADJ as
	'well yeah, it('s), it('s) like'

6	*se< ranta mei-l on semmone mu:rheen, (.)*
	DEM shore 1PL-ADE be+3SG DEM-ADJ trouble-GEN
	'that shore place we have is always trouble'.

In (7), EK makes an assessment about MS's summer house (line 1). JS responds to this with another assessment, bringing up negative aspects (lines 4–6). The negative assessment is prefaced with the particle *nii*. According to Sorjonen (2001b: 185), *nii* can be used as a weak claim of agreement in responses to affiliation-relevant utterances such as assessments, as in (7). Here, *nii* prefaces *em mä tiä* 'I don't know' which expresses epistemic stance and is followed by a downgrading assessment of the summer house (lines 5–6). *Em mä tiä* 'I don't know' is produced with accelerated tempo (marked with angle brackets > <) and the main verb appears in the phonologically reduced form *tiä* rather than the full form *tiedä*.

Similar to (7), *tietää* overwhelmingly occurs in clauses with negative polarity; 68% of the occurrences of *tietää* are in negative clauses (188/278). If we only consider *tietää* in the 1st person singular, the percentage of clauses with negative polarity is 82% (145/177). In the first person, the majority of clauses with *tietää* occur with pronominal subjects (87%; 154/177), just like in example (7). Our findings receive further support from a study on Estonian interaction by Keevallik (2003), who found a close correlation between 1st person subjects and verbs of cognition (such as *teadma* 'know' and *arvata* 'think').

The typical context for the verb *ajatella* 'think' is illustrated in (8):

(8) (SaPu119)

1 P3:   *ky:l se   aika pitkälti on,   kyl se*
            PTC DEM quite mainly be+3SG PTC DEM
            'it surely more or less is,'

2   *siis   jotku kirjaime-t lausu-taa   eri   tava-l*
             PTC some letter-PL   pronounce-PSS different way-ADE
             'surely some letters are pronounced differently'

3   *et   ässä lausu- eiku <u>see</u> lausutaa   niinku, useemma-l   tava-l.*
             COMP letter.S   PTC DEM pronounce-PSS PTC   various-ADE way-ADE
             'the letter "s" is pron- I mean it is pronounced in several different ways'.

4 P1:   *aha,*
             PTC
             'Oh'

5 P3:   *se   vaihtele-e.*
             DEM vary-3SG
             'It varies.'

6 P1:   *mu-l   o   semne setä jonka   nimi ol-i   Aleksi*
             1SG-ADE   be+3SG such   uncle REL+GEN   name be-PST+3SG MaleName
             'I had an uncle whose name was Aleksi'

7   *ja   sit   hän   ol-i   Arkenttiina-s   asu-i-vat   ni,*
             and   then 3SG   be-PST-3SG Argentina-INE live-PST-3PL   PTC
             'and then he was in Argentina (they) lived,'

8   *ni sit   ku   hä   tuli   sielt*
             so then   when 3SG   come-PST+3SG from.there
             'and when he came back from there'

9   *ni si-tä   sanot-t-i   Aleho-ks.*
             so 3SG-PAR   say-PSS-PST MaleName-TRA
             'he was called Alejo.'

10   *ni m(ä) aattel-i-n   et   on-k   on-k-s   se   niinku*
             so 1SG   think-PST-1SG COMP   be+3SG-Q be+3SG-Q-CLT 3SG   like
             'So I thought, is that, is that it'

11   *et   <Aleksi>   lausut-ta-is   Aleho.*
             COMP   MaleName pronounce-PSS-CON   MaleName
             'that Aleksi was pronounced as Alejo.'

12 P3:   *se   v<u>oi</u>   olla.*
             3SG may+3SG be-INF
             'It may well be.'

12 P1:   *°joo°*
              PTC
              'Yeah.'

In (8) the co-participants are discussing the pronunciation of Finnish names in a Spanish-speaking community. P1 is talking about her uncle (lines 6–9), and in lines 10–11 she is presenting an inference based on the facts she has related. The inference is formulated as a question (*onk se niin et Aleksi lausuttais Aleho* 'is it the case that Aleksi was pronounced as Alejo', and it is prefaced by *m(ä) aattelin et* 'I thought that', marking it as her personal inference. *M(ä) aattelin et* is phonologically eroded: The first two syllables of the finite verb are merged into one (*ajattelin > aattelin*) and the subject pronoun is merely a nasal sound *m*. The nasal sound of the pronoun forms a syllable together with the long vowel at the beginning of the finite verb.

     It is important to note that in (8), *m(ä) aattelin et* occurs in a context where the actions or activities of the speaker are not under discussion; the topic is the uncle. In line 6 the speaker introduces a new referent, the uncle, by anchoring the referent to herself. The uncle then becomes a new discourse topic. When drawing an inference based on what she has told the coparticipants about the uncle (lines 10–11), the speaker adds an epistemic stance marker *m(ä) aattelin et* (cf. Kärkkäinen 2012 on English *I thought*). The referent of *mä* 'I' does not become topical.

     The next example, (9), illustrates the use of the verb *luulla* 'think, believe' and *muistaa* 'remember'. The participants are discussing a deceased person and his heirs.

(9) (C133)

1 EK: *si-l   on   useampi tytär*
              DEM-ADE be+3SG more   daughter
              'he has several daughters'

2   *[>oli-k se< neljä   muistaak[seni.*
              was-Q   it four   remember-INF-1SGPX
              'four was it if I remember right'

3 LP: *[jaaha,*
              PTC
              'I see.'

4 MK: *[neljä mä   luule-n   [et   si-l   on.*
                                           four 1SG think-1SG COMP 3SG-ADEbe-3SG
                                           'four I think he has.'

5 LP: *[ai*
                                                                PTC
                                                                'Oh'

6   *jaha.*
              PTC
              'I see'.

7 EK: *ja   ei-ks   ne   oo   kaikki naimisissa.*
           and NEG+3SG-Q DEM.PL be.CONNEG all   married
           'and they're all married aren't they'

8   (0.5)

9 MK: *yks on   ainakin: jossain (.) Rahikkala-ssa   talo-emäntä-nä*
           one be+3SG at.least   somewhere Place.Name-INE farm-wife-ESS
           'one of them is at least somewhere in Rahikkala, she's a farmwife'

10   *n' m' >muista-n<.*
             1SG remember-1SG
           'I remember'

11   (0.6)

12 EK: *joo ja yks on   Pyölinranna-ssa ja,*
           PTC and one be+3SG Place.Name-INE and
           'yeah and one is in Pyölinranta and'

13 MK: *kyl niit   on   niit   tyttär-i-i.*
           PTC DEM.PL.PAR be+3SG DEM.PL.PAR daughter-PL-PAR
           'there are plenty of daughters.'

In line 2 EK makes a suggestion regarding the number of daughters of the deceased. The suggestion is formulated as a question seeking confirmation, is followed by an epistemic phrase *muistaakseni*; this is morphologically an infinitival form of the verb *muistaa* 'remember', with a first person possessive suffix which syntactically marks the subject of the infinitive. However, it has become crystallized, and in the recent comprehensive grammar of Finnish, it is characterized as an adverb expressing epistemic stance (Hakulinen et al. 2004: 1524), and as such is not an example of double-marking. With *muistaakseni*, EK expresses an epistemic stance towards the suggestion she has just made. MK responds to this by confirming the number of daughters (line 4).

The confirmation (line 4) unfolds in an interesting way syntactically: MK first gives the number (of daughters) with just the numeral in the nominative case. This is followed by *mä luulen et* 'I think that' which expresses reservations concerning the exact number she has just given. The verb *luulla* is a complement-taking predicate, and it is followed by a complementizer *et* and a complement clause. The complement is formed as a possessive clause construction [X_{Ade} + is + Y], where X expresses the possessor and Y the possessed. This construction takes the numeral as its argument. Thus the utterance-initial phrase is now interpreted as a member of the complement clause. Another way to view this would be to say that *mä luulen et* functions as an epistemic parenthetical or formulaic fragment (cf. Thompson 2002) or a prefab, "a prefabricated expression" (Erman & Warren 2000) which is inserted in the middle of the clause. However, it does have the morphosyntactic components of a construction consisting of a pronominal subject and a finite verb showing agreement with the subject. As Bybee (2010: 36) points out, "To say something is a prefab does not mean that it does not have internal structure". Instead, there are associations between the prefab and the more general construction from which it has developed – in this case

between the expression *mä luulen et* and a construction consisting of a subject and a finite verb agreeing with the subject.

Example (9) continues with EK's question concerning the daughters (line 7), seeking confirmation as to whether the daughters are all married. After a pause, MK starts answering with what looks like the beginning of a list of the whereabouts of the daughters (line 9): MK reports on one of the daughters. This is followed by an epistemic expression formed with the verb *muistaa* 'remember' in the first person singular form. The subject pronoun is merely a clitic pronoun attached to the verb. However, even if the subject is reduced in form, its pronominal form is still evident in the clitic pronoun. Thus, it is in contrast with the epistemic phrase *muistaakseni* in line 2 which, as we saw in our earlier discussion, has been lexicalized as an adverb.

The topic of the conversation in (9) is the deceased and his daughters. The referent of 'I' is brought in to express the speaker's stance towards the issues under discussion, but it does not become the topic of conversation. The same is true of examples (7) and (8): The construction containing the first person singular pronominal subject expresses the speaker's stance, but the topic of the conversation remains the same – the summer house and the uncle. This turns out to be typical of pronominal subjects in the singular 1st person with cognitive verbs.

Lindström et al. (2009) have studied the use of 1st person verbal and pronominal person markers in Estonian dialect data; using dialect interviews, they found that referential distance was a statistically significant factor affecting the appearance of the pronominal person marker. They analyzed the data with respect to referential continuity (cf. Givón 1983); it turned out that even though the first person refers to a speech act participant, which as such is given in the discourse context, the expression of the 1st person subject with a separate pronoun is dependent on whether or not the 1st person referent has been referred to in the preceding clause. The statistical analysis carried out by Helasvuo and Kyröläinen (2011) shows that in Finnish conversational data the majority of 1st and 2nd person pronominal subjects show a referential distance of 3 (or more), while 1st and 2nd person zero subjects favor a referential distance of 1. This means that for zero subjects the referent has typically been mentioned in the immediately preceding clause, for example in the first part of a coordinated clause or in the first pair-part of a question–answer adjacency pair (cf. section 0 above). In contrast, for pronominal subjects there are usually at least two or more clauses intervening between the pronominal subject and the previous mention of the same referent.

To consider these quantitative findings in the light of a closer analysis of the local contexts of the pronominal subjects: Our analysis of examples (7)–(9) has shown that 1st and 2nd person pronominal subjects typically occur in contexts where the topic of the conversation concerns something else, and the construction containing the 1st or 2nd person subject and a verb of cognition is brought in to express the speaker's stance towards the issue under discussion. The referent of the 1st or 2nd person subject does not become topical.

We have also seen that 1st and 2nd person subjects typically occur with complement-taking predicates (CTPs), but the syntactic relationship between predicate and complement is not always clear: There may not be any obvious complement, as in (7), or there is no complementizer, as in (9). The construction containing the 1st or 2nd person subject and a verb of cognition is often phonologically reduced and eroded, as in (7)–(9). These constructions form prosodic chunks (Bybee 2010). In particular example

(9) illustrates that these constructions function as chunks, which can be added either at the end of the utterance or in between.

Based on an analysis of complementation in English conversation, Thompson (2002) challenges the traditional view of complements as subordinate clauses that stand in a grammatical relation with a CTP. She suggests that they are better understood in terms of formulaic fragments expressing speaker stance toward the content of a clause. She also notes the connection between epistemic CTPs and 1st person subjects and proposes that the primary function of CTPs is to frame a clause in subjective epistemic terms. In her data these formulaic CTPs are most often epistemic, but they may also be evidential or evaluative. The most frequent formulaic CTPs involve the verbs *think*, *guess*, *remember*, and *know* (Thompson 2002: 138). Interestingly, this list is almost identical with our list of the most frequent verbs of cognition in the Finnish data (cf. Table 3). In Finnish, these verbs strongly favor the 1st person singular form, and pronominal subjects rather than zero subjects.

In her article on subject omission in Finnish, Duvallon (2006) suggests that one home environment for zero subjects in the 1st and 2nd person is crystallized expressions like *en tiiä* '(I) don't know'. The present data strongly imply, however, that this suggestion is mistaken. Duvallon's article focuses on subject omission only, and her data consist of examples of zero subjects; the contexts of pronominal subjects are not considered at all. Our conversational data show that crystallized expressions are far more common with pronominal subjects than with zero subjects. Constructions containing 1st or 2nd person subject pronouns and verbs of cognition might be seen as 'prefabs', prefabricated expressions (Helasvuo 2014). It is important to note that even though prefabs can be thought of as ready-made chunks, this does not mean that they do not have any internal structure. On the contrary, there are associations between the prefab and other occurrences of words that appear in the prefab, as well as between the prefab and the more general construction from which it has arisen. (Cf. Bybee 2010: 36.)

To sum up our discussion of the home environments of pronominal subjects: We can say that notions such as referential distance or verb type can be used as quantitative measures which help us find patternings in a large data set, but in order to explain and gain a deeper understanding of these patternings we need qualitative microanalysis of the local contexts in which the pronominal subjects occur. Our analysis shows that even though speech act participants are present in the speech situation and as such given, the occurrence of pronominal subjects is sensitive to previous mention: Pronominal subjects are preferred over zeros in contexts where the referent has not been mentioned for a while and the topic of the conversation is something else. The pronominal subject, even if it is expressed only as a clitic pronoun attached to the verb, allows for early projection of the trajectory of the turn. This is in contrast with the home environments for subject omission (section 3) where early projection is achieved either through the complex syntactic unit or sequence (e.g. in question–answer adjacency pairs through the first pair part) or the actions (e.g. in directives).

5. Questions as a special environment for subject expression

I now turn to one final environment for subject expression, namely questions. Duvallon (2006: 212) mentions questions as one further context in which 1st and 2nd person subjects are typically omitted. Questions are indeed an interesting context with regard to subject expression, but it is mistaken to link them straightforwardly with zero subjects. As I show in this section, patterns of subject expression in questions are considerably more complex.

Based on the current database, omission of the subject pronoun does not appear to be particularly frequent in questions. The database contains 109 utterances which function as questions and are formed either with the question clitic –*k(o/ö)* or with a question word such as *mitä* 'what', and the predicate verb is in the 2nd person singular. Only twelve, i.e. 11%, appear without a pronominal subject. This is less than the overall proportion of zero subjects in the 2nd person singular (cf. Table 1). It has been suggested that the omission of subject pronouns in questions is due to contextual cues such as politeness; the subject is more likely to be omitted if the speaker seeks to appear polite and not imposing (Lappalainen 2004: 197–199). This may explain the omission of the subject pronoun in a couple of examples in the data; these examples are all derived from one of the recordings included in the database, and were all produced by a waitress serving her customers.

In Finnish, polar questions are formed with a question clitic *-ko/-kö* attached to the finite verb. The verb appears clause-initially and is followed by the subject. Word order in polar questions is thus VS rather than SV, which is the typical word order in transitive and intransitive clauses (see Helasvuo 2001b: 75–81). In colloquial speech it is quite common in 2nd person constructions for the question clitic to be left out or shortened to just the initial consonant–*k*. If the question clitic is left out entirely, the utterance is marked as a question only through the reversed (VS) word order, as in (10):

(10) (D113a)

1 Sari:	*Oo-t*	*sä*	*käy-ny*	*jo*	*katto-o*	*si-tä.*
	be-2SG	2SG	go-PCP	already	see-INF	3SG-PAR
	'Have you already been to see it?'					

2 Satu:	*E-n,*	*mu-l*	*ei*	*o-*
	NEG-1SG	1SG-ADE	NEG+3SG	have-CONNEG
	'No, I don't have - -'			

In (10) Sari asks Satu a polar question (line 1). There is no question clitic; rather, the utterance is marked as a question with the VS word order. With SV order the utterance would be a declarative (*sä oot käyny kattoo* 'you have been to see [it]'). In cases like (10) it is of course not possible to leave out the subject pronoun because the utterance on line 1 is marked as a question only through the reversed word order: If there was no subject there would be no possibility of reversing the order.

Sometimes the shortened form of the clitic appears as –*ks*, where the final –*s* can be interpreted either as a clitic particle or possibly as a cliticized form of the 2nd person singular pronoun *s(in)ä* ('you (sg)'). However, as I will show later (in connection with ex. (12)), the latter option seems less likely. Consider ex. (11).

(11) (D113b)

1 Mikko: *aio-k-s   muute   kopsa-ta   se-n   diablo[peli-n*
               plan-Q-CLT   by.the.way   copy-INF   it-ACC   diablo.game-ACC
               'By the way, do you plan to copy Diablo'

2 Santeri: *[minkälaise-n*
                                             what.kind.of-ACC
                                           'what kind of'

3   *kossu-n   sä   meinaa-t   homma-ta   sinne*
               hard.drive-ACC 2SG   plan-2SG   get-INF   there
               'hard drive do you plan to get there'

4   (0.2)

5 Ben: *yks piste kuus giga-ase-n (.)   avde aflon äkspe-n. (.) vitut.*
               one dot six   giga-ADJ-ACC Name-ACC   shit
               'A 1.6 G Avde Aflon XP [name of the type of hard drive]. Shit.'

6   *(.) maksa-a   kakssataaviiskytä euroo. (.) hä mutta mä hommaa-n (.)*
                 cost-3SG 250   euro-PAR   but   1SG get-1SG
               '(It) costs 250 euros. But I will get'

7   *jonku-n (.) yks piste kakkose-n tai yks piste nelose-n ˚sinne˚.*
               some-ACC one dot two-ACC or one dot four-ACC there
               'some 1.2 or 1.4 [different types of hard drives] there.'

8 Santeri: *no   joo*
               PTC PTC
               'oh well'

9   (.)

10 Mikko: *nii aio-k-s   kopsa-ta velje-lle-s   diablo-n.*
               PTC plan-Q-CLT copy-INF brother-ALL-2SGPX   diablo-ACC
               'So do you plan to copy Diablo for your brother?'

11 Ben: *ei   se* ((CLEARS THROAT))   *halun-nu   si-tä.*
             NEG+3SG   3SG   want-PCP   3SG-PAR
             'He didn't want it.'

In line 1, Mikko asks a question which is directed to Ben. The last two syllables of Mikko's turn are overlapped by Santeri, who asks another question. Ben first responds to Santeri's question; Santeri's brief receipt token (line 8) is followed by a short pause (line 9), offering Ben an opportunity to come back to Mikko's question in line 1. Ben does not take this opportunity, and Mikko repeats his question in line 10. Rather than responding to the question directly, Ben explains in line 11 why he does not intend to do what Mikko's question has referred to.

Both of Mikko's questions, in lines 1 and 10, are formed with the question clitic -*k*, attached to the finite verb. The question clitic is followed by another clitic -*s*, which appears in directives and questions, in both wh-questions and polar questions (Hakulinen et al. 2004: 801). Its meaning has been described as conveying familiarity (L. Hakulinen 1979: 88–89). The fused question clitic –*k(o)s* is said to refer to (assumed) common knowledge (Hakulinen et al. 2004: 801). The latter explanation, however, does not seem to describe its use in (11) lines 1 and 5. If familiarity is interpreted as the casual nature of the question, this may explain the choice here of the fused clitic –*ks*.

Interestingly, in connection with the fused question clitic -*ks* the verbs in lines 1 and 10 have no 2nd person singular suffix (-*t*). Thus these verbs have no person marking at all, by means of either the personal pronoun or the verbal person suffix. Rather than a matter of omission, however, the 2nd person verbal suffix should be seen as assimilated with the -*k* of the question clitic (Hakulinen et al. 2004: 69). The assimilation is motivated by the fact that -*tks* is not a possible consonant cluster in Finnish. The form is recognized as a 2nd person singular form that stands in contrast with other person forms (see (12) below). It is important to note that the clitic -*s* can be used in connection with questions and directives not only with the 2nd person singular but with all personal forms, as in (12):

(12) (C158a)

Mirja:	*keitä-n-k-s*		*mää*	*tee-tä.*
	boil-1SG-Q-CLT		1SG	tea-PAR
	'Shall I make some tea?'			

In (12), the fused clitic –*ks* is attached to the finite verb. The verb carries a 1st person singular suffix –*n*, and there is also a 1st person singular subject pronoun *mää* 'I'. Thus, even though the clitic –*s* can be analyzed as a clitic originating in the 2nd person singular pronoun, it no longer functions as a personal form, standing in a paradigmatic relationship with other personal forms. This means that in (11), the questions in lines 1 and 10 are not doubly marked.

In sum, questions form a special environment for subject expression. Contrary to what has been suggested in the literature (Duvallon 2006: 212), subject omission is not particularly frequent in questions. We also noted that there may be some issues concerning politeness that could work in favor of subject omission in certain contexts, such as customer service, but the current database does not provide a suitable basis for a deeper investigation of this. We further found that polar questions with no question clitic and a 2nd person singular subject, as in (10), do not allow the possibility of zero subject. Finally, examples (11) and (12) show that the question clitic –*ko/kö* often appears in a shortened form –*k*, which is sometimes followed by another clitic –*s*. While this clitic -*s* has been shown to originate in the second person singular pronoun, example (12) shows that it no longer functions as a person marker.

6. Conclusions

I have discussed the home environments for zero and pronominal subjects in the 1st and 2nd person singular. In order to find motivations for the choice of the form of subject, I have analyzed them in terms of the micro-level social actions performed by the participants. Based on a syntactically coded database of conversational Finnish, I have shown that in both 1st and 2nd person, there is a clear preference for pronominal subjects over zeros; in other words, double-marking is preferred over single-marking. This clearly contravenes the general preference for minimization or economy in reference to person in conversation, as suggested by Sacks and Schegloff (1979) and Levinson (2007; see also Hacohen and Schegloff 2006).

Sacks and Schegloff (1979: 16) also identify a general preference for recipient design, according to which speakers use reference forms which allow the recipients to recognize who is being referred to (cf. also Levinson (2007) on the principle of recognition). As we have seen in the analysis of the Finnish data, there are various phonological processes that result in assimilation and erosion, affecting both pronominal and verbal person markers (cf. especially sections 4 and 5). It is possible that double marking functions to ensure recognition despite possible assimilation or erosion.

We may also note that the presence of a pronominal person marker may have an effect on projection. Verbal person markers are suffixal in Finnish; because of the preference for SV order in declaratives as discussed in section 5, verbal person markers come later in the turn than do possible pronominal markers. In terms of projection, the expression of the subject with a pronoun allows for early projection of the future trajectory of the turn. In the home environments for subject omission, however, early projection can be achieved through other means. As we have seen in section 3, many of the environments involve complex structural units. When the construction with the omitted subject is embedded in a complex structural unit, such as a question–answer adjacency pair, same-subject coordination, or self-repair targeted at the verb, there is a pronominal subject in the larger sequence that allows early projection by the co-participants, even when the pronominal subject is lacking in the construction itself.

The database used in this article has been systematically coded for several morphosyntactic and discourse features that may affect the choice of the form of subject expression. One purpose of the study has been to show how we can gain important insights regarding the global patterns that arise from a systematic analysis of the data. However, we cannot understand the deeper motivations that lie behind these patternings without a close microanalysis of the local contexts of subject expression. The database may serve as a basis for recognizing recurrent patterns and building collections of examples sharing certain features.

Acknowledgements

Appendix

Abbreviations used in the glosses

1SG	1st person singular pronoun or verbal suffix
1SGPX	1st person singular possessive suffix
2SG	2nd person singular pronoun or verbal suffix
2SGPX	2nd person singular possessive suffix
3SG	3rd person singular pronoun or verbal suffix
3PL	3rd person plural verbal suffix
ACC	accusative
ADE	adessive
ALL	allative
CLT	clitic particle
CMP	comparative
COMP	complementizer
CON	conditional mood
CONNEG	connegative verb form
DEM	demonstrative
ESS	essive
GEN	genitive
ILL	illative
INE	inessive
INF	infinitive
NEG	negation verb
PAR	partitive
PCP	participle
PSS	passive
PST	past tense
PTC	particle
Q	question clitic
REL	relativizer
TRA	translative

References

Ariel, Mira (1990) *Accessing noun-phrase antecedents*. London: Routledge.

Bybee, Joan L. (2010) *Language, usage, and cognition*. Cambridge: Cambridge University Press.

Clayman, Steven E. (2012) Turn-constructional units and the transition-relevance place. In Jack Sidnell, and Tanya Stivers (eds.), *The Handbook of Conversation Analysis*. Chichester: Wiley-Blackwell, pp. 150-166.

Erman, Britt, and Beatrice Warren (2000) The idiom principle and the open choice principle. *Text* 20.1: 29–62.

Dixon, R.M.W. (2005) *A New Approach to English Grammar on Semantic Principles*. 2nd edition. Oxford: Oxford University Press.

Dryer, Matthew S. (2011) Expression of pronominal subjects. In Matthew S. Dryer, and Martin Haspelmath (eds.), *The World Atlas of Language Structures Online*. Munich: Max Planck Digital Library, chapter 101. Available online: http://wals.info/chapter/101. Accessed on 2012-04-02.

Duvallon, Outi, and Antoine Chalvin (2004) La réalisation zéro du pronom sujet de première et de deuxième personne du singulier en finnois et en estonien parlés. *Linguistica Uralica* XL.4: 270–286.

Duvallon, Outi (2006) Milloin pronominisubjekti jää pois puhutussa suomessa? In A. Pajunen, and H. Tommola (eds.), *XXXII Kielitieteen päivät Tampereella 19.–20.5.2005*. Tampere: Tampere University Press, pp. 203-217.

Givón, Talmy (1983) *Topic continuity in discourse: A quantitative cross-language study*. Amsterdam: John Benjamins Publishing Company.

Hacohen, Gonen, and Emanuel A. Schegloff (2006) On the preference for minimization in referring to persons: Evidence from Hebrew conversation. *Journal of Pragmatics* 38: 1305–1312.

Hakulinen, Auli, and Lea Laitinen (2008) Anaforinen nolla: Kielioppia ja affekteja. *Virittäjä* 112: 162–185.

Hakulinen, Auli, Maria Vilkuna, Riitta Korhonen, Vesa Koivisto, Tarja Riitta Heinonen, and Irja Alho (2004) *Iso suomen kielioppi*. SKST 950. Helsinki: Finnish Literature Society.

Hakulinen, Lauri (1979) *Suomen kielen rakenne ja kehitys*. 4th edition. Helsinki: Otava.

Helasvuo, Marja-Liisa (2001a) Emerging syntax for interaction: Noun phrases and clauses as a syntactic resource for interaction. In Elizabeth Couper-Kuhlen, and Margret Selting (eds.), *Studies in Interactional linguistics*. Amsterdam: John Benjamins Publishing Company, pp. 25-50.

Helasvuo, Marja-Liisa (2001b) *Syntax in the making: The emergence of syntactic units in Finnish conversational discourse*. Studies in Discourse and Grammar 9. Amsterdam: John Benjamins Publishing Company.

Helasvuo, Marja-Liisa (2014) Agreement or crystallization: Patterns of 1st and 2nd person subjects and verbs of cognition in Finnish conversational interaction. *Journal of Pragmatics* 63: 63-78.

Helasvuo, Marja-Liisa, and Lea Laitinen (2006) Person in Finnish: Paradigmatic and syntagmatic relations in interaction. In Marja-Liisa Helasvuo, and Lyle Campbell (eds.), *Grammar from the human perspective: Case, space and person in Finnish*. CILT 277. Amsterdam: John Benjamins Publishing Company, pp. 173-207.

Keevallik, Leelo (2003*) From interaction to grammar. Estonian finite verb forms in conversation*. Studia Uralica Upsaliensia 34. Uppsala: Acta Universitatis Upsaliensis.

Kibrik, Andrej A. (2011) *Reference in Discourse*. Oxford: Oxford University Press.

Kärkkäinen, Elise (2003) *Epistemic Stance in English Conversation. A Description of Its Interactional Functions, with a Focus on* I think. Amsterdam: John Benjamins Publishing Company.

Kärkkäinen, Elise (2007) The role of I guess in conversational stancetaking. In R. Englebretson (ed.), *Stancetaking in Discourse: Subjectivity, evaluation, interaction.* Amsterdam: John Benjamins Publishing Company, pp. 183-219.

Kärkkäinen, Elise (2012) I thought it was very interesting. Conversational formats for taking a stance. *Journal of Pragmatics* 44: 2194-2210.

König, Ekkehard, and Peter Siemund (2007) Speech act distinctions in grammar. In Timothy Shopen (ed.), *Language Typology and Syntactic Description*. Cambridge: Cambridge University Press, pp. 276-324.

Lappalainen, Hanna (2004) *Variaatio ja sen funktiot. Erään sosiaalisen verkoston jäsenten kielellisen variaation ja vuorovaikutuksen tarkastelua*. Helsinki: Finnish Literature Society.

Lappalainen, Hanna (2006) Pronominisubjektin käytöstä ja poisjätöstä Kelan asiointikeskusteluissa. In Taru Nordlund, Tiina Onikki-Rantajääskö, and Toni Suutari (eds.), *Kohtauspaikkana kieli. Näkökulmia persoonaan, muutoksiin ja valintoihin*. Helsinki: Finnish Literature Society, pp. 37-64.

Levinson, Stephen C. (2004) Deixis. In Laurence R. Horn, and Gregory Ward (eds.), *The Handbook of Pragmatics*. Oxford: Blackwell, pp. 97-121.

Levinson, Stephen C. (2007) Optimizing person reference – perspectives from usage on Rossel Island. In N.J. Enfield, and Tanya Stivers (eds.), *Person reference in interaction: Linguistic, cultural and social perspectives*. Cambridge: Cambridge University Press, pp. 29-72.

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

Oh, Sun-Young (2005) English zero anaphora as an interactional resource. *Research on language and social interaction* 38: 267–302.

Oh, Sun-Young (2006) English zero anaphora as an interactional resource II. *Discourse studies* 8.6: 817–846.

Östman, Jan-Ola (1981) *You know: A discourse-functional approach*. Amsterdam: John Benjamins Publishing Company.

Sacks, Harvey, and Emanuel A. Schegloff (1979) Two preferences in the organization of reference to persons in conversation and their interaction. In George Psathas (ed.), *Everyday Language: Studies in Ethnomethodology*. New York: Irvington, pp. 15-21.

Scheibman, Joanne (2002) *Point of View and Grammar: Structural Patterns of Subjectivity in American English Conversation*. Studies in Discourse and Grammar 11. Amsterdam: John Benjamins Publishing Company.

Siewierska, A. (1999) From anaphoric pronoun to grammatical agreement marker: Why objects don't make it. *Folia Linguistica* 33.1–2: 225–251.

Sorjonen, Marja-Leena (2001a) Simple answers to polar questions: The case of Finnish. In Margret

Selting, and Elizabeth Couper-Kuhlen (eds.), *Studies in interactional linguistics*. Amsterdam: John Benjamins Publishing Company, pp. 405-431.

Sorjonen, Marja-Leena (2001b) *Responding in conversation.* Amsterdam: John Benjamins Publishing Company.

Sulkala, Helena, and Merja Karjalainen (1992) *Finnish. Descriptive Grammars*. New York: Routledge.

Tao, Hongyin (1996) *Units in Mandarin Conversation: Prosody, Discourse, and Grammar*. Amsterdam: John Benjamins Publishing Company.

Tao, Hongyin (2001) Discovering the usual with corpora: The case of remember. In Rita Simpson, and John Swales (eds.), *Corpus Linguistics in North America: Selections from the 1999 Symposium*. Ann Arbor: University of Michigan Press, pp. 116-144.

Thompson, Sandra A. (2002) 'Object complements' and conversation: Towards a realistic account. *Studies in Language* 26.1: 125–164.

Thompson, Sandra, and A. Anthony Mulac (1991) A quantitative perspective on the grammaticization of epistemic parentheticals in English. In Elizabeth Traugott, and Bernd Heine (eds.), *Grammaticalization* II. Amsterdam: John Benjamins Publishing Company, pp. 313-339.

Torres Cacoullos, Rena, and Catherine Travis (2011) Testing convergence via code-switching: Priming and the structure of variable subject expression. *International Journal of Bilingualism* 15: 241–267.

Weber, Elizabeth G., and Paola Bentivoglio (1991) Verbs of cognition in spoken Spanish: A discourse profile. In Suzanne Fleischman, and Linda Waugh (eds.), *Discourse pragmatics and the verb: Evidence from Romance*. London: Routledge, pp. 194-213.