


ARTICLE

The Finnish tail construction as a first mention

Katri Priiki 

School of Languages and Translation Studies, FI-20014 University of Turku, Finland
Email for correspondence: katri.priiki@utu.fi

(Received 22 May 2019; revised 4 May 2020; accepted 18 May 2020; first published online 22 September 2020)

Abstract

This article examines the Finnish tail construction (right dislocation) used as a first mention of a referent and the variation of the demonstrative pronouns *tämä* ‘this’, *tu* ‘that’, and *se* ‘it’ in the construction. Many previous studies have suggested that tail construction (TC) referents are highly active and thus already mentioned and salient in a conversation. However, in Finnish, the TC may introduce new referents into a conversation, and this article provides an empirical analysis of how and why this is done. First-mention TCs are often evaluations or questions in which the proposition links the utterance to the preceding context. When presenting new information, the TC allows the speaker to present a potentially lengthy lexical definition of a new referent at the end of the utterance, avoiding the additional emphatic meanings or unwanted implications a simply inverted word order might create.

Keywords demonstratives; dislocation; Finnish; tail; word order

1. Introduction

Spoken language exhibits a group of linguistic structures in which a referent is referred to twice within the same utterance: one reference, usually a pronoun, occurs in the typical place of the argument, and another reference, often a lexical noun phrase, occurs in an atypical place. Often, the lexical noun phrase is placed either before the canonical clause structure, as in (1a), or after, as in (1b); the noun phrase is presented in italics and the pronoun in bold.

- (1) a. *Your friend Mary*, I saw **her** yesterday.
b. I saw **her** yesterday, *your friend Mary*.

These structures are often called DISLOCATIONS (e.g. Geluykens 1987, 1992) or DETACHMENTS (e.g. Amon 2015), implying that the noun phrase has been moved from the place where it ostensibly belongs and replaced with a pronoun. Thus, (1a) would exemplify LEFT DISLOCATION and (1b), RIGHT DISLOCATION. These terms have been criticized because no actual movement is involved and because dislocation has a connotation that supports ‘a view that spoken language is a defective form of written language’, as Timmis (2009: 330) put it. Moreover, describing the extra elements

Table 1. Distribution of pronouns in TCs of the Satakunta corpus.

Pronoun used	All	First mentions
<i>se</i> ‘it’ + <i>ne</i> ‘they’	171 + 21	67 + 4
<i>tämä</i> ‘this’ + <i>nämä</i> ‘these’	26 + 3	16 + 1
<i>tuo</i> ‘that’ + <i>nuo</i> ‘those’	23 + 1	20 + 1
Mixed	18 + 3	8 + 1

as ‘right’ or ‘left’ refers to the written form, obscuring the fact that these structures are typically phenomena of spoken language.¹

In this article, I examine constructions where the extra element is presented after the canonical clause structure, exemplified by (1b), and I use the terms TAIL and TAIL CONSTRUCTION (see Carter & McCarthy 1995, McCarthy & Carter 1997). With tail construction (TC), I refer to the entire structure, which comprises both the CORE CLAUSE and the TAIL NOUN PHRASE (NP). By CORE CLAUSE, I mean the canonical clause structure without the tail NP. I call the pronoun occurring in the core clause of the TC the PLACEHOLDER PRONOUN in order to distinguish it from pronouns used as determiners of the tail NP. In Finnish, the same demonstrative pronoun is typically used as both the placeholder pronoun and the determiner of the tail NP, as (2a–c) show (see Etelämäki 2006: 68² and Table 1 in Section 4 below for the distribution of different variants in the data of this article). In examples (2a–c) and later, the placeholder pronoun is highlighted in bold and the tail NP in italics.³

- (2) a. **mitä** *se* *teki se sulhanen* *sitte* (Sapu190⁴)
what it did it bridegroom then
 ‘What did he do then, the bridegroom?’
- b. **tämä** *ei miu’u ollenka tämä kissa* (Arkisyn, D131)
this NEG meow at.all this cat
 ‘This (one) does not meow at all, this cat.’
- c. **toi** *hoitaa sitä toi Martta nyt* (Sapu188)
that takes.care it.PART that NAME now
 ‘That (woman) takes care of her now, Martta.’

TCs are traditionally associated with marking of topic, and with reference to English data, it has been claimed that tail constituents are always highly salient; that is, the information status of the element referred to twice in the TC must be discourse-old (Lambrecht 1994: 202; Ward, Birner & Huddleston 2002). Some researchers have explained the TC as self-repair, adding the lexical NP as an after-thought to disambiguate or clarify the referent when the speaker notices that a pronoun is insufficient (Geluykens 1987: 122). Some TCs may well be clarifications the speaker decides to produce mid-utterance, but some may be premeditated choices.

The view that the TC is always used to maintain a topic (e.g. Lambrecht 1994: 181) has also been challenged: in spoken French, the TC is used at least occasionally to present new information (Ashby 1988: 213; De Cat 2007: 526) and to change topic (Pekarek Doehler, de Stefani & Horlacher 2015: 133–135). By TOPIC I mean simply what is talked about in a turn (Maynard 1980: 263), which does not always correspond to the THEME of a clause (see Section 3). According to Etelämäki (2006: 66–67), a speaker may express with her choice of word order what she offers as a topic, but the participants decide together whether they continue to discuss the topic or not.

Pekarek Doehler et al. (2015: 2–3) suggested that the TC is related to the information structure of the utterance, allowing a speaker to modify the place of an element in turn. I elaborate on this suggestion by examining data on the Finnish first-mention TC. My results, presented in Section 4, support the notion that the TC often occurs as an evaluation (e.g. Aijmer 1989, McCarthy & Carter 1997, Timmis 2009). By EVALUATION I mean that the clause expresses the speaker's subjective opinion, perception or stance.

I show that in Finnish, it is possible to introduce new referents with the TC, exploiting the presenting and pointing functions of the demonstrative pronouns. The TC offers a useful opportunity to present a lexical description of a new referent at the end of an utterance, while the placeholder pronoun of the structure gives a cue as to how the hearer should interpret the referent and also maintains the neutral word order. Examining a language with rich morphology and flexible word order (e.g. Vilkuna 1989) will contribute new insights to the research on the TC. I continue Karhu's (1994) and Etelämäki's (2006) work on contemporary conversational data.⁵ The available corpora are large enough for some quantitative observations, focusing on pronoun variation in the core clause of the construction and on the information structure of the first-mention TCs. Etelämäki's (2006) study focuses on the pronoun *tämä* 'this' and Karhu's (1994) data mainly include variants with *se* 'it'. However, the Finnish TC variant with the pronoun *tuon* 'that' in (2c) above has been studied very little, and my analysis provides new information by comparing it to other variants.

This article is structured as follows. In Section 2, I discuss previous research on the TC and its relation to word order. Section 3 introduces Finnish demonstrative pronouns used in the TC. My data and numerical information about them are described in Section 4. Section 5 offers an analysis of the TC as a first mention of a referent, and Section 6 briefly concludes the study.

2. The tail construction and information structure

The TC is viewed as a common feature across numerous languages (Dik 1978: 140), yet languages exhibit significant differences with respect to how the TC is defined grammatically and the frequency of its use. In spoken English, TCs are relatively rare (Winkle 2015), but their frequency is sensitive to the type of conversation: they are usually found in casual and unplanned multiparty conversations involving evaluative comments. Timmis (2009) showed that in English, the TC is a consistent and durable feature and can be found even in a corpus of 70-year-old conversations.

Previous studies (e.g. Lambrecht 1994, Pekarek Doehler et al. 2015 and, specifically on Finnish, Karhu 1994) have shown that the functions of the TC have a connection with managing topics. However, linguists disagree about whether the TC maintains and continues a topic (Lambrecht 1994: 183) or signals a topic change (Pekarek Doehler et al. 2015). Karhu (1994) suggested that in Finnish dialect interviews, the TC is used to restructure conversation and change topics. She also showed that the placeholder pronoun in the TC connects the utterance to the preceding context, while the tail NP ensures that the utterance is understandable as an independent structure. Etelämäki (2006) continued this line of argument by suggesting that in the TC, the persistence and information status of a referent may conflict: a referent may be both new and presented as persistent, or it may be both briefly mentioned and familiar.

I expand the study of the Finnish TC by approaching occurrences that are used to introduce new topics from the viewpoint of word order and information structure. The term INFORMATION STRUCTURE was first introduced by Halliday (1967). Lambrecht (1994: 5) defined the study of information structure as examining (i) the propositional information and the presuppositions and assertions, (ii) the identifiability and activation of referents, and (iii) what the proposition⁶ is about and what the topic of an utterance is.

Finnish imposes very few grammatical constraints on the order of the constituents⁷ in a clause. Instead, the word order is discourse conditioned (Vilkuna 1989: 9–16). In Finnish grammar, the verb initial field is divided into two elements: the THEME (in Vilkuna's terminology, the T FIELD) is presented immediately before the finite verb, while before the theme, a topicalized element (the K FIELD) may be added. New information, the RHEME,⁸ is presented at the end of the utterance, in the predicate verb and after it. Vilkuna (1989: 73) suggested that 'a markedly late position in a sentence may be a good way to introduce a brand-new entity in the discourse'. The most neutral order of the constituents is subject–verb–object, and the relevant distinctions are marked by case inflection (Vilkuna 1989: 21, 42–43). Even though the different variations of the word order do not affect the interpretation of a subject or an object, the variations carry certain emphatic meanings that make them usable only when such interpretation is desired. Moreover, in certain clause types, such as possessive and existential clauses, the order of constituents is quite fixed.

In this article, I only examine a subset of TCs that is defined more strictly than in many previous studies. Namely, I only focus on structures consisting of a pronominal constituent as a placeholder in the core clause and a co-referential tail that is a lexical NP or a pronoun. In my data, NP tails are far more frequent than pronoun tails, which are marginal. My data also contain TCs in which the placeholder is a demonstrative adjective form, such as *semmoinen* 'like that', and the tailed element describes a quality, as seen in (3a) below. These are excluded because in these cases, defining a first mention is more problematic than with pronouns and locative adverbs. When, as in (3b), the placeholder pronoun *se* 'it' is in the partitive, *sitä*, it sometimes behaves like a pragmatic particle so that the co-referential tail is hard to define (Vilkuna 1989: 139–145). Karhu (1994: 22) also viewed utterances such as (3c), in which the pronoun is repeated immediately, as TCs. I have also excluded types (3b, c) from my data.

- (3) a. **Semmoiseksihan** se muuttuu *arkipäiväiseksi*. (Vilkuna 1989:141)
such.TRA.CL it changes plain.TRA
 ‘That’s what it becomes, plain.’
- b. Kai **sitä** sentään koti pitää ihmisellä olla. (Vilkuna 1989:139)
I.guess it.PART after.all home must person.AD be.INF
 ‘Surely a person should, after all, have a home.’
- c. Kyllä **se se mies** menee sinne taloon. (Karhu 1994:22)
Yes it it man goes there house.ILL
 ‘Surely the man goes into the house.’

In French and Spanish, TCs do not always have a pronoun or a clitic placeholder (De Cat 2007: 509). Focusing on a narrow subset of TCs in this article is practical for examining a language with flexible word order, such as Finnish. In other words, if the placeholder pronoun is not used, the NP situated in the end of the clause will often be interpreted as a regular clausal constituent, not a tail NP, within the possible variations of word order (see Amon 2015: 149 for similar problems with Estonian).

Some studies (e.g. Geluykens 1987: 122; Helasvuo 2001) have viewed the TC as resulting from self-repair. It has also been proposed (Fretheim 1995, Ziv & Grosz 1998) that it would be possible to distinguish repairs and genuine TCs by examining, for example, intonation contour and the position of the tail NP. However, this suggestion is based on invented examples, and I agree with Amon’s (2015) conclusion that in genuine speech, repairs and TCs form a continuum without reliable criteria with which to draw a clear distinction. In this article, I concentrate on occurrences of the TC that do not include a substantial number of hesitation markers, such as pauses and repetitions, and that are thus unlikely to be word searches or other processing problems in the first place.

Pekarek Doehler et al. (2015) suggested that the TC is a means of dealing with problems of reciprocity: the TC is formed when the speaker produces a new ending for their utterance, creating a new opportunity for the recipient to react. In my data, as (2) above illustrates, TCs are usually produced as a single intonation unit, without prosodic features that would signal the ending of the turn between the core clause and the tail NP.⁹ Moreover, in Finnish, the tail NP is not always the final constituent of an utterance – recall (2a, c). When there is a clear pause, a response from another speaker or an ending intonation contour before the lexical NP, I define the NP as an independent NP, not as a tail NP.

Relevant earlier studies on the TC have approached the subject from different viewpoints by applying the frameworks of either conversation analysis (e.g. Etelämäki 2006) or information structuring (e.g. Amon 2015). To be able to make comparisons between my observations based on my data and earlier observations based on other data, I try to take into account both interactional and grammatical phenomena when analysing the examples in this article.¹⁰ The concepts and the terminology used to study word order are diverse and even confusing, and I try to tackle this problem by explicitly defining the terms I use.

Because this article examined the TC as a first mention, it is necessary to define first mention. When I refer to a FIRST MENTION or a NEW referent, I mean LOCALLY INITIAL OCCASIONS, which is what Schegloff (1996: 450) called references that have no preceding mentions in the conversation. As locally initial forms, lexical references are preferred, and a pronoun – alone or as a part of the TC – is a clearly marked choice. A participant's ACCESSIBILITY to a referent, in turn, indicates whether the hearer can identify the referent. Identification may be based on a previous mention; the context or an evoked frame; or the participants' shared common ground (e.g. Vilkkuna 1992: 16). Identification that is enough in a certain conversational context may not always be precise knowledge of the referent or identical interpretation of it for all participants (Du Bois 1980: 232–258). Sometimes, sufficient identification is understanding the role of the referent in the action the speaker does with her turn (Etelämäki 2006: 16–17).

In the framework of conversation analysis, reflecting a speaker's assumptions is avoided, but in functional approaches to grammar, the choice of referential form is seen to reflect the speaker's understanding of whether or not other participants know the referent. In this article, I mostly follow the functional tradition in this sense, but when analysing the examples, I also consider relevant conversation analytic studies and take note of participants' reactions. Following Lambrecht (1994: 76–77), I define as ACTIVE those referents that a speaker presents as easy for the hearer to identify. In Schegloff's (1996: 451) description, this definition corresponds to occasions when a speaker treats a referent as a continuation of the earlier conversation. With ACCESSIBLE referents, sufficient identification is possible even if it requires more effort from the hearer than with active referents, and INACTIVE referents are not accessible as such (Lambrecht 1994: 76–77). In Du Bois' (1980: 221–226) definitions, accessible referents correspond to cases where the reference 'requires a hearer to look elsewhere to interpret them' and inactive referents are NONIDENTIFIABLE, wherein 'identification with some known referent is not possible'.

3. Finnish demonstratives in the tail construction

Finnish has three demonstrative pronouns: *tämä* 'this', *tuo* 'that', and *se* 'it'. In spoken Finnish, *tämä* is usually realized as *tää* and *tuo* as *toi*. These colloquial forms are originally Southern, but today they are so widely distributed that they do not carry a strong regional connotation. *Tuo* 'that' has another colloquial variant, *tua*, which is limited to Southwestern dialects and sometimes occurs in the data of this study.

All the demonstrative pronouns may be used either independently or adnominally, and, in colloquial speech, they are often used to refer to people even though Standard Finnish has a separate third person personal pronoun, *hän* 'he, she'. In colloquial style, however, *hän* is used mostly logophorically, referring to a speaker of a reported utterance (see Laitinen 2005, Priiki 2017b), and the neutral anaphoric way of referring to people is *se* 'it'.

The various discourse functions of Finnish demonstrative pronouns have been widely studied using an interactional approach and in naturally occurring data

(Laury 1997; Seppänen 1998; Etelämäki 2006, 2009; Priiki 2017a). Traditionally, the pronouns have been viewed as expressing the proximity of the speaker or the hearer to the referent (Larjavaara 1990), but studies of conversational data have shown that factors crucial to the interpretation of a demonstrative reference are cognitive, social, and affective rather than spatial (e.g. Östman 1995, Laury 1997), and that the pronouns may be used interchangeably (e.g. Etelämäki 2006: 102–110). The spatial distinctions activate, however, when two or more referents are contrasted (Reile et al. 2019). I agree with Laury's (1997) and Etelämäki's (2006, 2009) view that by using different pronouns, speakers focus attention on referents and express their perception of the accessibility of a referent, which is connected to the ongoing activity.

Etelämäki (2006, 2009) described Finnish demonstratives according to their referentiality and indexicality, which are dimensions presented by Hanks (1992). Indexically, demonstratives imply either symmetrical (*se* 'it', *tuo* 'that') or asymmetrical (*tämä* 'this') ground (Etelämäki 2006, 2009), or, in Laury's (1997) terms, they imply that the referent is in the speaker's sphere of attention (*tämä*), the hearer's sphere (*se*), or outside both (*tuo*). In terms of accessibility, this means that a referent referred to with *se* or *tuo* is presented as equally active (or inactive) for both the speaker and the hearer. When a reference is made with *tämä*, the speaker has primary access, and thus the situation is asymmetrical. In the scale of referentiality, the pronouns *tämä* and *tuo* are open; that is, in an utterance in which *tämä* or *tuo* is used, directing attention to the referent is relevant because the referent is still being identified (Etelämäki 2006, 2009).

As shown above in (2), all Finnish demonstratives – and occasionally even the third person personal pronoun *hän* – may occur in the TC and as the first mention of the referent. In general, a third person or demonstrative pronoun as a locally initial form is a rare phenomenon because, in the presentation of new referents, lexical descriptions are preferred (Schegloff 1996: 450). In my previous study (Priiki 2017a:58), I found that only 6% of these pronouns are used as first mentions. This makes the TC as a first mention a clearly marked choice.

In another previous study (Priiki 2015), I argued that the relative frequencies of pronouns in TCs are different from those in canonical clauses. The pronouns referring to people in the third person singular are distributed in conversational data as follows: the vast majority (83%) are occurrences of the anaphoric demonstrative pronoun *se* 'it'; next in frequency (14%) is the logophoric pronoun *hän* 'he, she'; and occurrences of *tämä* 'this' and *tuo* 'that' are rare (both around 2%). As the placeholder demonstrative of the TC, *se* 'it' is still the most frequent, with a share of 75%. The percentages of the two other demonstratives and *hän* are, however, opposite: in TCs, 12% of pronoun phrases are occurrences of *tuo* 'that', 11% are occurrences of *tämä* 'this', and only 3% are occurrences of *hän* 'he, she' (Priiki 2015).

With new referents, the distribution of pronouns resembles the one in TCs: the standard personal pronoun *hän* 'he, she' is marginal (around 1%), and *tämä* 'this' and *tuo* 'that' are more frequent (*tämä* 10%, *tuo* 12%) than in the data in general. In fact, these two phenomena intertwine: the TC placeholder pronouns are quite often first mentions of a referent, and the relative frequency of the pronouns *tämä* and *tuo*, referring to a person, is highest when the pronoun is both a TC placeholder and a first mention of a referent (*tämä* 14%, *tuo* 21%; Priiki 2015).

This means that in contrast to English, in which ‘tailed’ referents are previously mentioned or otherwise highly active, in spoken Finnish, the TC is regularly used to introduce new referents. Particularly if the placeholder demonstrative of the TC is *tämä* ‘this’ or *tuo* ‘that’, the referent may not be mentioned previously. In particular, a reference performed with *tämä* ‘this’ may imply the primary access of the speaker. A reference with *tuo* ‘that’, on the other hand, may treat the referent as equally inaccessible (or irrelevant) for both the speaker and the hearer. In Section 5 below, I examine these kinds of occurrences further.

4. The data of the study

The study’s data consist of two collections of conversations. The first comprises 24 hours of contemporary multiparty conversations (face-to-face) with 33 speakers that were audio-recorded in Southwestern Finland in 2009–2012 during the University of Turku’s research project ‘Satakunta in Speech’.¹¹ I refer to this dataset as the Satakunta corpus. From the transcribed conversations, I manually collected the total of 266 instances of the TC. Of these, 117 (44%) present a referent that is not mentioned in the preceding recorded conversation at all, or at least in the preceding 10 or more turns.¹² The distribution of pronouns used in TCs in the data is presented in Table 1. In addition to the forms listed in Table 1, the data contain 12 TCs in which a locative adverb form (*siellä*, *siäl* ‘there’, *tääl*, *täsä* ‘here’, or *tual*, *tosa* ‘over there’) is used as a placeholder.

The majority of the TCs in the Satakunta corpus are formed with the demonstrative pronoun *se* ‘it’, and plural pronouns are much less frequent than singular pronouns. Variants with bare NP tails, as exemplified in (8) below, are classified according to the placeholder pronoun. There are 29 occurrences¹³ of bare NP tails in the data, and 17 of them are proper names, such as the one shown in (8). The variants, in which different pronouns are used as the placeholder and the determiner pronoun of the tail NP, are presented as ‘mixed’ in Table 1. The sum of mixed variants also includes four instances in which the personal pronoun *hän* ‘he, she, (it)’ is used as a placeholder in the TC. In the Satakunta corpus, 34% of the TCs express evaluation and 12% are questions; among the first-mention TCs, these respective numbers are 37% and 13%. Since there are no data available on the average frequencies of questions and evaluations in conversational Finnish, it is hard to tell how significant these numbers are. In spoken English, according to Siemund’s (2018: 378) corpus search, around 8% of clauses are interrogative.

To supplement the analysis of the rarest variants, I have collected another dataset of 35 TCs with *tämä* ‘this’ and 35 with *tuo* ‘that’, together with another six, mixed constructions. This dataset was constructed by manually selecting TCs from the results of an automated search¹⁴ of Arkisyn, the morphosyntactically coded corpus of conversational Finnish (approximately 30 hours of recordings and 327,000 words). Arkisyn includes video-recorded, face-to-face conversations and audio-recorded, telephone conversations. In the Arkisyn collection, 21 (60%) TCs with *tämä* ‘this’ and 21 TCs with *tuo* ‘that’ present a previously unmentioned referent. Of these 42 first-mention TCs, 22 are evaluations and eight are questions. See Appendix A for additional details of both corpora.

Table 2. The continuity of the TC referents in the data. First-mention placeholder pronouns and the number of times their referents are mentioned again. (Satakunta corpus + Arkisyn collection).

Placeholder pronoun	Mentions of the referent		
	0-1	2-4	5 or more
<i>se</i> 'it' / <i>ne</i> 'they'	32	21	22
<i>tämä</i> 'this' / <i>nämä</i> 'these'	7 + 9	6 + 8	4 + 4
<i>tuo</i> 'that' / <i>nuo</i> 'those'	10 + 16	6 + 4	4 + 1

5. The tail construction as first mention in the data

5.1 Variants of the tail construction with different demonstrative pronouns

The TC with the anaphoric demonstrative pronoun *se* 'it' is the most frequent and most examined variant of the Finnish TC. In Karhu's (1994: 430) dialect interview data, almost all TCs have *se* 'it' or its plural variant *ne* 'they' as the placeholder pronoun. Karhu suggested that the motivation for using the TC is that the placeholder pronoun presents the referent as known and the tail NP verifies the correct interpretation. Etelämäki's (2006) view is that the placeholder pronoun indexes how the current utterance relates to the preceding conversation. According to Etelämäki, by using a pronoun phrase for introducing the referent instead of a lexical NP, a speaker implies that the referent is meant to be continuous; that is, it would repeatedly be referred to in the conversation (see Du Bois 1980).

In this section, I argue that implying discourse continuity may not be the main reason for choosing a first-mention TC. In my data, the majority of first-mention TCs introduce a referent that is referred to again only once or not at all. Table 2 presents how many times the referent that the TCs introduces is referred to in turns that follow the first-mention TC.

I will first introduce typical occurrences of the first-mention TCs with *se* 'it' and *tämä* 'this' in my data to compare them with the variant with *tuo* 'that'. In (4), the speakers are teenagers and young adults. The transcription symbols used in (4) and later are explained in Appendix B.

- (4) 1 Jere: ooks sää Satu nähny's sen Kummeli alivuokralaise,
'Have you, Satu, seen the (film) "Kummeli alivuokralainen"?"
- 2 Satu: oo. eiks me oltu Sakke kattoos [leffas.
'(I) have. We saw (it) in the movie (theater), didn't we, Sakke?'
- 3 Sakke: [joo.
- 4 viis sekunttia Afrikas (ni voi vittu), <hetkine>
'Yeah. Five seconds in Africa so, "Oh fuck, (wait) a moment".'
(0.8)
- 5 Jere: vähä **se** oli kämä:ne *se* (.) *se* yks kohta.
little it was lame it it one section
'It was so lame, the scene.'
- 6 hei hei hei mä oon Steeffan. mut (.) Stifu: jengille.
'Hi, hi, hi, I'm Stefan. But Stifu to the gang.'

(Sapu154)

On the first line, Jere introduces a new topic by asking his older cousin Satu whether she has seen a certain Finnish film. The name of the film is marked definite by the demonstrative pronoun *se* ‘it’ as a determiner (see Laury 1997). Satu gives a positive answer, and the boys begin to recall scenes from the film.

In line 5, Jere refers to a scene of the film with a TC, using the pronoun *se* ‘it’ as both the placeholder and the determiner of the tail NP. The scene to which he refers has not been mentioned in the conversation, but the mention is interpretable in the context of the film, which, after the question in the beginning of the excerpt, is explicitly known to all participants. Jere’s turn has some indications of word search: the pronoun *se* is repeated, and there is a slight pause before the tail NP. The noun *kohta* is quite vague: it means ‘point’, ‘spot’, or ‘section’, and a more precise description, which allows the hearer to identify the scene, is a direct quote.

The tail NP *se yks kohta* has two determiner pronouns: the demonstrative *se* ‘it’ and *yks* ‘one’. Their combination implies that shared knowledge of the referent is assumed, but interpreting it may require more information from the speaker (Hakulinen et al. 2004: Section 1417). The reference is a prospective indexical (see Goodwin 1996), prefacing the description (i.e. the quote) and inviting the hearers to recognize the scene.

The TC enables Jere to present his evaluation of the scene – that it was *kämäne* ‘lame’ – before the rather lengthy quote. Thus, the hearers already know Jere’s stance to the scene when they recognize the scene. This gives them an opportunity to align (or disalign; see Goodwin & Goodwin 1987: 14) early on in the statement. The other participants align with a short laugh after the quote. The scene mentioned in (4) is not a continuous referent. After the excerpt, Jere repeats his evaluation *vähä se oli kämäne* ‘it was so lame’ and after this, the participants move on to evaluate other scenes of the movie.

In (5), the participants, a mother (Seija) and her daughter (Kati), are talking about Kati’s childhood friends, who are presumably known to Seija as well.

- (5) 1 Seija: Hanna on aika ussein siä kans.
 ‘Hanna is quite often there [on Facebook], too.’
 2 Kati: joo: se on aika [ahk(era).
 ‘Yeah, she is quite active.’
 3 Seija: [nin o joo. (1.5)
 ‘She is.’
 4 .hh no mitäs **sen** elämään kuuluu *sen* (.) *Aneten* sitte.
 well what it.GEN life.ILL belongs it.GEN NAME.GEN then
 ‘Well, what’s going on in her life, Anette’s, then?’
 (lit.: ‘What belongs to her life?’)
 5 Kati: ↑no e::: sil on kans kradu- hh
 ‘Well, no- she too has thesis -’
 6 Seija: [hommelit viä.
 ‘Stuff still.’
 7 Kati: [kradunteko tässä
 ‘Thesis work now.’

(Sapu187)

The person introduced into the conversation with a TC using the pronoun *se* ‘it’ (Anette, line 4) is not mentioned in the recorded conversation before Seija’s turn. In the beginning of the excerpt, Seija mentions Hanna, who is another childhood friend of Kati and whose news they have been discussing.

The TC that introduces Anette (line 4) is used to change the topic from Hanna to another one of Kati’s childhood friends. The ongoing action¹⁵ is changing as Seija moves from evaluating Hanna’s Facebook behaviour to asking about Anette. With the particle *sitte* ‘then’ in the end of her turn, she contrasts Anette to Hanna (see Hakulinen et al. 2004:Section 825), thus making the move explicit. However, updating old friends’ news is the main line of the conversation since before Hanna and Anette, Seija and Kati have already discussed other people known to both. Presenting the proposition (having something new in life) before the name of the new referent allows Seija to express early in her turn that she is returning to asking about news. The placeholder pronoun *se* ‘it’ implies that the previously unmentioned referent will be identifiable to the hearer. In contrast to (4) above, the referent in (5) becomes a continuous topic: Anette’s situation is discussed further in many turns that follow.

While the interpretation of a first-mention *se* ‘it’ is based on shared knowledge, the frame for interpreting a reference with a first-mention *tämä* ‘this’ is the current utterance (Laury 1997, Etelämäki 2006). In other words, a turn with *tämä* ‘this’ includes some information that may help the hearer understand the reference or link the turn to the previous conversation. According to Etelämäki (2006: 49), a first mention with *tämä* ‘this’ is often in the subject position, and the predicate of the clause presents a new definition of the subject. Laury (1997), on the other hand, suggested that a first mention with *tämä* ‘this’ signals that the referent is accessible to the speaker but not necessarily to all other participants.

In my data, topics are opened and activated using a TC with *tämä* ‘this’ when a speaker wants to propose a topic on which he or she is focused but when other participants are focused on other things. In (6), the same two speakers as in (5) are visiting friends.

- (6) 1 Seija: mitäs tarskos meitin ruveta hh hinnaamaan itteemme johonkin päi.
 ‘What (about)- Should we start moving (lit. towing ourselves) somewhere?’
- 2 Kati: mh.
- 3 Seija: **tää** on aina tämmöstä kellon kytäämistä
this is always like.this.PART clock.GEN watching.PART
- 4 [sit *tää läh-* mt
then this
 ‘This is always this kind of clock watching, this (leaving).’
- 5 Raili: [niinku sit täytyy lähtii-
 ‘Like then (you) have to leave -’
- 6 Merja: [nih
 ‘Yeah.’

(Sapu 202)

The excerpt begins with Seija proposing that they should consider leaving by using a humorous choice of words.¹⁶ She then justifies the need to leave with a turn that is

formulated as a TC with the pronoun *tämä* ‘this’. The tail part is left unfinished, but the syllable *läh* projects the word *lähtö* or *lähteminen*, which both mean ‘leaving’. Seija has previously stated that she will have a night shift at work, so they need to leave early enough.

The TC with *tämä* ‘this’ in (6) functions similarly to those examined by Etelämäki (2006: 75–78). In her turn, Seija defines the referent of *tämä* ‘this’ and, in my interpretation, gives a general evaluation of what leaving is like when you have a timetable (recall the word *aina* ‘always’), thus redefining the departure of Seija and Kati, which was introduced into the conversation in line 1. Moreover, Seija expresses her own point of view and her personal experience: the verb she uses, *kytätä* ‘watch, stalk, lurk’, has a strong negative connotation. If the pronoun was *se* ‘it’ or *tuu* ‘that’, the turn would create an assumption that the other participants had independent access to the phenomena she describes. In (6), the continuity of the reference is hard to define. The concept of ‘leaving’ is not referred to again in the conversation, although they continue to talk about the practical details of leaving.

The TC in (6) is atypical among the data in the sense that the other participants begin their responses before Seija has uttered the tail NP. There is no final intonation contour or pause before the tail. In my data, when there is overlap in a TC, the responses usually begin between the determiner pronoun and the noun of the tail NP. The early onset is explained so that, by responding in overlap (see Vatanen 2014), Raili and Merja demonstrate that they recognize the situation Seija is describing – even though Seija’s turn does not imply an expectation that they will. Again, by using the TC, Seija is able to describe her experience before explicitly naming the concept she is talking about. As the example shows, the other participants explicitly express that they know what Seija means even before she has produced the lexical description.

The speakers in (7) below are high school students doing their homework and other tasks together. Before the excerpt, they are each focused on their own books and papers. Lotta is copying a list of contact details for a hobby they all share.

- (7) 1 Lotta: kukas **tää** o *tää* *Elina*.
 who this is this NAME
 ‘Who is this (girl), this Elina?’
 (0.3)
- 2 Milja: .h se on mikä oli se, (.) *semmone*,
 ‘She is- what was it, (someone) like,’
- 3 Lotta: °a [i nii°.
 ‘Oh’
- 4 Milja: [muuttanu sielt jostai. .hh
 ‘(Who) moved from somewhere.’

(Arkisyn, sg120)

Formed as a TC, Lotta’s question (line 1) picks up a name from the paper in front of her. Thus, the referent is active for Lotta but not for the others, and with her choice of demonstrative pronoun, she guides the others to locate the referent in the list on which she is focusing. There is no pointing gesture involved: the pronoun together

with Lotta's gaze is adequate for locating the referent. The reference is easily identified by the recipient: in her answer, Milja refers to Elina with the pronoun *se* and thus treats the referent as active and known enough. Elina is not mentioned in the recorded conversation before the excerpt, and after the question–answer sequence shown, she is not discussed further. Thus, as already shown in Table 2, even TCs with *tämä* 'this' do not always introduce continuous referents, as was argued by Etelämäki (2006).

First-mention TCs with the pronoun *tuo* 'that' are rare in Etelämäki's (2006) data. Similarly, in Laury's (2005a) study of 'first and only' pronoun references, she encountered only one *tuo* pronoun. However, in my dataset that is larger than Etelämäki's and Laury's, there are several occurrences of *tuo*, as presented in Section 4. Almost all TCs with *tuo* 'that' or the plural *nuo* 'those' in my data refer to relatively unimportant referents, by which I mean that the participants would be able to continue the conversation even though their interpretation of the referent was vague (for interpretation of referents, see Vilkuna 1992: 16–25; Etelämäki 2006: 16–17).

When *tuo* 'that' is used as a first mention, the focus of the utterance is on the proposition of the clause and action that the speaker does with it. In my data, the referent introduced into a conversation with *tuo* and a tail NP does not usually remain as a topic for long (see Table 2). In the beginning of (8), participants agree that a microwave oven is an essential kitchen appliance nowadays.

- (8) 1 Kari: kyä se taitaa k- melkein °kaikilla ollaj jo°.
'(I think) almost everyone already has it.'
- 2 Eeva: hjoo.
'Yeah.'
- 3 Pia: kyl se o- melkein ku- (.) melkein kuuluu £sivistyks(h)ee(h)
4 [j(h)o£ h h
'It is- almost belongs to the culture already.'
- 5 Kari: [jaa **tolla** ei muuten oo *Ranella*.
oh that.AD NEG otherwise be NAME.AD
'Oh, that (guy) doesn't have one, Rane(, for that matter).'
- 6 Pia: nii si:llä ei oom mut [se: om muutenki [semmone,
'Yes, he doesn't, but he's like that anyway.'
- 7 Kari: [ei o. [se sano et siittä
8 tullee noita semmosia sätteitä
'(He) doesn't. He said that those- some rays come out of it.'
- (Sapu188)

Kari mentions his acquaintance, Rane, who does not have a microwave oven (line 5). Rane has not previously been mentioned in the recorded conversation. The clause is a possessive clause, which in Finnish presents the owner in the adessive case (ending *lla*). The turn begins with the particle *jaa*, which is often used for receiving news (Hakulinen et al. 2004: Section 798) and also to mark remembering and noticing (*Kielitoimiston sanakirja* (2018), s.v. *jaa*). Kari uses it to mark his turn as expressing something that just came up in his mind. This is enforced with the adverb *muuten*, which may be translated as 'otherwise', 'by the way', or 'for that matter'. The pronoun *tuo* 'that' points out a referent that is not the centre of attention for either the speaker or other participants (Priiki 2017a).

My interpretation is that the main content of Kari's utterance is to point out an exception to the view he presented earlier, that is, to deny the proposition of having a microwave oven being true for everyone rather than to introduce Rane as a topic. By using a TC, Kari is able to present the proposition, which is the only link between Rane and the preceding conversation, before mentioning Rane by name. The action Kari makes with his turn does not require other participants to actually even know Rane. Piia, however, does, and Piia and Kari continue discussing Rane for a few turns. After that, the discussion shifts to suspicions people used to have about microwaves. Creating a contrast is another feature of the pronoun *tuo* 'that' used in (8) (see Laury 1997: 103–104; Seppänen 1998: 208): it distances Rane from microwave oven owners and thus also from Kari and his collocutors.

In (9), a family is discussing the fact that university students often have to work unpaid trainee jobs as a part of their studies. The excerpt begins with Mari trying to remember information she has previously read about the number of unpaid working hours.

- (9) 1 Mari: ja se oli siis jotai ihan niinkun, .hhh hhh £tyylii
 2 k(h)uuskyyttuhatta t(h)ai j(h)otain n(h)iinku, k(h)uussataatuhatta£=
 3 =ei nyt ehkä >niim paljom mut siis< iha hirveesti, .hhh. (0.7)
 'And it was something even like- like sixty thousand- something like
 sixty hundred thousand- maybe not so much, but I mean really a lot.'
 4 Jussi: joo-o, (0.3) **toi** on tommonen *toi yliopistomaailma*
yeah that is like.that that university.world
 5 että se on vähä niinku, (0.5) siin joutuu elelee semmosessa
 @kompromissien
 6 välimaastossa@ koko ajan.
 'Yeah, it [lit.: that] is like that, that university setting, that it is a little
 like- there (you) have to live kind of "in the midst of compromises"
 all the time.'

(Arkisyn, sg441)

The TC with *tuo* 'that' picks up a referent and a description from the preceding conversation even though the noun *yliopistomaailma* 'university setting' has not been lexicalized in the recorded conversation before Jussi's turn. The demonstrative adjective *tommonen* 'like that' refers to the description ('you have to do unpaid work'), and Jussi's turn expresses that he agrees with the thoughts just presented. In (9), the turn with the TC defines the university setting by presenting the mentioned practice as an established convention. At the same time, the reference using *tuo* implies contrast and distance between the speaker and the referent, as in the previous example (8): Jussi dissociates himself from the practice described and examines the university setting from afar.

In the Satakunta corpus, four of the 21 first-mention TCs with *tuo* 'that' have a tail that consists only of a pronoun, and the lexical definition of the referent is left incomplete. In these cases, my interpretation is that the choice of pronoun implies that the speaker has trouble accessing the definition of the referent. By choosing the pronoun *tuo* 'that', a speaker implies equal access – in these cases, that the referent is equally inaccessible to the speaker and the hearer. In (10), the speaker is an old

woman referring to a person she met during her visit to a health care centre, which she has previously discussed in the same conversation.

- (10) 1 *nyk ko toi (.) phh (.) toikin tuala otti toi mikä (.)*
now when that that.CL over.there took that what
- 2 *toi (.) ö- (.) toi noin ni se (.) mikä tualla oli niin (.)*
that that so well it that over.there was so
- 3 *sano et @sulla onkin kauheet käsivoimat@*
said that you.AD is.CL terrible.PL arm.power.PL
 ‘Now when that (one) there, also took- that (one) who- erm well, she/he who was there, said that, “You have such powerful arms.”’

(Sapu187)

In (10), the speaker’s problem with the referring expression may be that she does not remember whether the person was a doctor, a nurse, or some other type of personnel. She refers to the person repeatedly with the colloquial form of the pronoun *tuo* ‘that’, and to the health care centre with the locative adverb *tuala* ‘over there’. Some of the references may be ambiguous as hesitation particles (for the connection of the pronoun and the hesitation particle *tota*, ‘well, erm’ see Etelämäki & Jaakola 2009). A lexical description for either of the referents in this context is not produced. On line 2, the speaker changes the person reference to *se* ‘it’, implicating that the description is accessible enough for this context. It is indeed accessible enough, given the fact that none of the other participants asks for clarification on the referents. The main content of the utterance is to present the quote about the strength of the speaker, ‘You have such powerful arms’, which the speaker produces after word search and hesitation. In the context of (10), only the quote is relevant, and a very vague identification of the original speaker is sufficient. Because (10) is apparently connected to word search and, thus, differs from the previous examples, I do not analyse it further in the following section.

5.2 Why the tail construction as a first mention?

I have argued that the TC is a means of presenting a proposition about a new referent before introducing the referent lexically. This gives the hearer an opportunity recognize the situation described before the lexical NP is produced and to align (or disalign) with the evaluation as soon as they recognize the referent. The proposition also connects the utterance to the preceding conversation and provides a reason why the new referent is introduced in the current turn. I reproduce the TCs, which I presented above in context (4)–(9), in (11). As I mentioned in Section 2 above, in Finnish grammar, the theme is presented immediately before the finite verb. Before a theme constituent, a topicalized element may occur. The constituent in the theme position is underlined in (11) and also in (12), below.

- (11) a. *vähä se oli kämä:ne se (.) se yks kohta.*
little it was lame it it one section
 ‘It was so lame, the scene.’

- b. no mitäs sen elämään kuuluu *sen* (.) *Aneten* sitte
well what it.GEN life.ILL belongs it.GEN NAME.GEN then
 ‘Well, what’s going on in her life, Anette’s, then?’
 (lit.: ‘What belongs to her life?’)
- c. tää on aina tämmöstä kellon kytäämistä sit *tää läh-*
this is always like.this.PART clock.GEN watching.PART then this
 ‘This is always this kind of clock watching, this (leaving).’
- d. kukas tää o *tää Elina.*
who this is this NAME
 ‘Who is this (girl), this Elina?’
- e. jaa tolla ei muuten oo *Ranella.*
oh that.AD NEG otherwise be NAME.AD
 ‘Oh, that (guy) doesn’t have (one), Rane, for that matter.’
- f. joo-o, (0.3) toi on tommonen *toi yliopistomaailma*
yeah that is like.that that university.world
 ‘Yeah, it [lit.: that] is like that, that university setting.’

In (11a), the new referent is not interpretable with the noun *kohta* ‘point’, ‘spot’, or ‘section’. As mentioned above, the reference is a prospective indexical that prefaces the description. The proposition (being lame) expresses the speaker’s stance and shows that he is evaluating the film. In (11b), the proposition (having something new in life) expresses that the turn continues asking for news about acquaintances. In (11c), the lexical description of the referent is left quite unclear, but the other participants recognize the situation from the proposition (being ‘clock watching’). In (11e), the person mentioned has no connection to the previous conversation, but the proposition that is denied (having a microwave oven) does.

I have analysed the connection of the proposition of first-mention TCs to the preceding conversation in the Arkisyn collection (21 examples with *tämä* ‘this’ and 21 with *tuo* ‘that’). In 12 cases of a first-mention TC with *tuo* and in seven cases with *tämä*, I would say that the proposition is more important than the referent for the flow of the conversation. I suggest that this is at least one of the reasons why speakers choose such an order of constituents instead of constructed neutral word order variants of the same utterances, as seen in (12) below.

The example (11d) is an exception, wherein the proposition does not have a link to previous conversation. Rather, in (11d), the reason for choosing a TC is the newness of the referent. This is examined by comparing the original utterances to the constructed variants presented in (12). Utterances in (12) have, in principle, a neutral word order, which, according to Vilkuna (1989: 21, 42–43), is subject–verb–object and carries little emphatic meaning compared to other variants. However, the information structure of the utterances in (12) is different from that in (11) and the implications (i.e. what is known and what is new) are different.

- (12) a. vähä se yks kohta oli kämäne
little it one section was lame
 ‘The scene/one scene was so lame.’

- b. no mitäs sen Aneten elämään kuuluu sitte
well what it.GEN NAME.GEN life.ILL belongs then
 ‘Well, what’s going on in Anette’s life, then?’
- c. tää lähteminen on aina tämmöstä kellon kyttäämistä
this leaving is always like.this.PART clock.GEN watching.PART
 ‘Leaving is always this kind of clock watching.’
- d. kukas tää Elina o
who this NAME is
 ‘Who is this Elina?’
- e. jaa tolla Ranella ei muuten oo
oh that.AD NAME.AD NEG otherwise be
 ‘Oh, that Rane doesn’t have (one), for that matter.’
- f. joo-o, (0.3) toi yliopistomaailma on tommonen
yeah that university.world is like.that
 ‘Yeah, the [lit.: that] university setting is like that.’

While in (11) theme slots were occupied by pronoun phrases (except (11b)), in (12), there are rather long lexical noun phrases in the theme slot. Lexical phrases are typical references for new referents, but new referents are not usually brought up in theme position. According to Lambrecht (1994), in the theme position, references to inactive referents are the least acceptable constituents, even though sentences would be syntactically well-formed as such. In Finnish as well, anaphoric pronouns, which usually refer to active referents, are the most prototypical theme constituents (Vilkuna 1989: 41). Etelämäki’s (2006: 56) view is that first-mention lexical NPs with pronoun determiners bring up non-continuous referents, while first mentions with pronoun phrases introduce continuous referents.

To sum up, utterances in (12) are possible, but they differ from those in (11) in typicality and possibly discourse function. As mentioned above, the theme constituents of (12) are quite complex with their determiners, especially (12b). Moreover, the original utterances in (11a) and (11c) have some marks of processing trouble, so one of the reasons for choosing a TC may be that doing so gives the speaker more time to find a suitable lexical description, as was observed in (10). Presenting a lexical NP that describes a new referent at the end of a turn is aligned with the tendencies in which, firstly, new information, and secondly, long and complex constituents are likely to occur at the end of an utterance rather than at the beginning (Quirk et al. 1972: 943; Leech 1983: 83). These two tendencies are distinct phenomena (END-FOCUS and END-WEIGHT), but they intertwine: new information is often introduced into conversation with complex phrases, while known information is preferably referred to with short phrases, such as pronouns.

In Finnish, it would be possible to place the new, complex constituent at the end of the utterance even without the TC, as in the constructed examples in (13).

- (13) a. vähä oli kämänen se yks kohta
little was lame it one section
 ‘The/one scene was so lame.’

- b. no mitäs kuuluu sen *Aneten* elämään sitte
well what belongs it.GEN NAME.GEN life.ILL then
 ‘Well, what’s going on in Anette’s life, then?’
- c. on aina tämmöstä kellon kyttäämistä tää lähteminen
is always like.this.PART clock.GEN watching.PART this leaving
 ‘Leaving is always this kind of clock watching.’
- d. kukas o tää *Elina*
who is this NAME
 ‘Who is this Elina?’
- e. jaa ei muuten oo tolla *Ranella*
oh NEG otherwise be that.AD NAME.AD
 ‘Oh, (it) is not at Rane’s (place), for that matter.’
- f. joo-o, (0.3) on *toi yliopistomaailma* tommonen
yeah is that university.world like.that
 ‘Yeah, the (lit. that) university setting is like that.’

Why is the word order not simply inverted? Why is a placeholder pronoun needed?

Comparing the utterances in (13) to those in (11) and (12) demonstrates that in (13), the constituent occurring before the finite verb – if there is one at all – is an adverb, a particle, or a question word. Even though almost any word order may occur in Finnish, there are restrictions due to the clause type. For instance, the Finnish possessive clause, such as (12e), does not allow much variation without changing the meaning from possession to temporary location (Hakulinen et al. 2004: Section 896). Thus, the reading of (13e) without context would not be that something is not in Rane’s permanent possession, but rather that something is not located at Rane’s place or is not being used temporarily by Rane. However, the context would decide which way it would be understood, and sometimes such word order as in (13e) is used to emphasize the role of the possessor (Hakulinen et al. 2004: Section 896). The questions (13b) and (13d) are atypical, because typically the question word occurs before the theme (see Hakulinen et al. 2004: Section 1682), although they are possible in some contexts. In these questions, the question word is in the theme position instead of the placeholder pronoun (11d) or the NP, of which the placeholder pronoun is part (11b). This affects the focus of the question (see *ibid.*).

Even in basic clause types, the different variations of word order usually carry different emphatic meanings. If the finite verb is the initial constituent in a clause in which the default order is the subject first, the utterance will be interpreted as an emphatic reaction to something said before; strong agreement or disagreement; complaining; or bemoaning (Hakulinen et al. 2004: Sections 834, 1386). Thus, the utterance (13c) clearly gives a more complaining impression than (11c) or (12c), and (13e) has an insistent overtone.

When the subject of a clause occurs after the finite verb form, it will be interpreted as new information – as it is in these cases – or its meaning may be indefinite (Hakulinen et al. 2004: Sections 1370, 1378). In (13a), the difference between (11a) and (12a) is not very clear, but in (13a) the indefinite interpretation (‘one scene’ instead of ‘the scene’) is preferred.

Thus, I propose that the first-mention TC has a special function of expressing information structure. In some cases, it may be the only functional means of structuring the utterance so that the lexical NP with new information is presented at the end of the utterance without violating the fixed order of a clause type or adding unwanted emphatic meaning.

As shown above, the TCs that introduce new referents are often evaluations or questions. In evaluations, it is useful to present the proposition before the lexical (and potentially lengthy) description of a new referent, because this order gives a hearer time to align or disalign (see also Pekarek Doehler et al. 2015: 145–147). When the TC presents a new referent, presenting a proposition before the lexical description may also help a hearer to link the new referent to the preceding conversation. In clause types such as possessive clauses and questions, in which the word order is more fixed, the TC can relate the message the speaker wants by avoiding emphatic meanings that the inverted word order would add. At the same time, the TC can serve the purpose of presenting new and complex constituents at the end of the utterance.

6. Conclusions

I have examined the Finnish TC used as a first mention of a referent in a conversation. I have also investigated the variation of the demonstrative pronouns *tämä* ‘this’, *tuo* ‘that’, and *se* ‘it’ used as the placeholder pronouns in the construction. Previous studies have suggested that TC referents are highly active and thus already mentioned and salient in a conversation. However, in Finnish, the TC, particularly the variants of the TC with the pronoun *tämä* or *tuo*, may introduce new referents into a conversation.

The Finnish TC has been studied using dialect interview data, in which the variation of pronouns used proved to be scarce (Karhu 1994), and using conversation data with a focus on the pronoun *tämä* ‘this’ (Etelämäki 2006). The present article completes the linguistic study of first-mention TCs and of the variants with the pronoun *tuo* ‘that’, which have not had many occurrences in the datasets of previous studies, with conversation data that is large enough to provide quantitative observations.

When presenting new information, the reason for choosing a TC is connected to word order and information structure. The TC allows the speaker to present a potentially lengthy lexical definition of the referent at the end of the utterance, where new information is typically presented. At the same time, it avoids conveying additional emphatic meanings or unwanted implications, which a simply inverted word order would create, because in the TC the placeholder pronoun occupies the typical theme slot.

In the data examined here, the first-mention TCs are often evaluations and questions. In these actions, the TC is a useful structure when the proposition presented in the clause links the utterance to the preceding context or is otherwise more salient than the referent. In evaluations, the hearer gets an early opportunity to align or disalign with the evaluation presented and in questions, the proposition may clarify how the question is linked to the preceding conversation.

Acknowledgements. I would like to thank the anonymous reviewers for their valuable comments and suggestions. Writing of the article was partly funded by the Alfred Kordelin Foundation and Eino Jutikkala Fund, for which I am grateful.

Notes

- 1 Sometimes, tail and topic constructions occur in written language. In these cases, according to Blasco-Dulbecco (1999: 94), they are used to convey special stylistic effects. About Estonian TCs in written texts, see Amon (2015: 87–93).
- 2 Etelämäki (2006: 68) does not specify the amount of the different variants of the TCs in her data.
- 3 Glossing abbreviations used in the examples: AD = adessive case, CL = clitic, GEN = genitive case, ILL = illative case, INF = infinitive, NEG = negation verb, PART = partitive case, PL = plural, TRA = translative case.
- 4 This and similar annotations indicate example sources. See Section 4 below for details.
- 5 The data examined in this article are significantly larger than Etelämäki's data (12 hours or recordings) and represent a different genre than Karhu's dialect interview data: Satakunta corpus includes approximately 24 hours of recordings and Arksyn corpus includes approximately 30 hours of recordings.
- 6 By PROPOSITION, I mean the lexical content of a clause; that is, what it means without considering its interactional function.
- 7 By CONSTITUENT, I mean a word or a group of words that functions as a single unit (e.g. in subject position).
- 8 The terminology used to discuss word order and information structure is not consistent. For example, Lambrecht (1994) used the terms TOPIC and COMMENT instead of THEME and RHEME.
- 9 In Finnish, the prosodic contour of a turn typically steadily falls towards the end, and a new turn begins with a higher intonation (Hakulinen et al. 2004: Section 1010).
- 10 The frameworks are not always easily combined, and there has not been a thorough discussion on their relation to each other. In this article my focus is on analysis on empirical data, not on forming a theoretical-methodological synthesis.
- 11 Satakunta is a region in Southwestern Finland.
- 12 In a previous study (Priiki 2015), I observed that checking 10 turns is sufficient for quantitative analysis in defining whether a reference is anaphoric. In the examples presented in this article, when necessary, I have inspected more thoroughly whether a referent is discussed in the recorded conversation more than 10 turns previously and whether it is treated as a new or recurring topic.
- 13 The distribution of placeholder pronouns with bare NP tails is as follows: 18 *se*, five *tuo*, two *ne*, two *hän*, one *tämä*, and one *nuo*.
- 14 I searched utterances with two demonstrative pronouns in one utterance. From the results, which were mostly not TCs, I picked up the first 35 occurrences of TCs with *tämä* and *tuo* and all mixed constructions I encountered so far. A disadvantage of this kind of search was that TCs with bare tail NPs could not be found.
- 15 By ACTION, I mean speech actions that are performed with each turn of talk, e.g. asking, answering, evaluating, aligning, etc.
- 16 Seija speaks quite strong dialect: *tarskos meitin* 'should we' would be *tarvitsisiko meidän* in standard Finnish.

References

- Aijmer, Karin. 1989. Themes and tails: The discourse functions of dislocated elements. *Nordic Journal of Linguistics* 12, 137–154.
- Amon, Marri. 2015. *Initial and Final Detachments in Spoken Estonian: A Study in the Framework of Information Structuring*. Ph.D. thesis, University of Tartu.
- Ashby, William J. 1988. The syntax, pragmatics, and sociolinguistics of left- and right-dislocations in French. *Lingua* 75, 203–229.
- Blasco-Dulbecco, Mylène. 1999. *Les dislocations en français contemporain. Etude syntaxique*. Paris: Champollion.

- Carter, Ronald & Michael McCarthy.** 1995. Grammar and the spoken language. *Applied Linguistics* 16(2), 141–158.
- De Cat, Cécile.** 2007. French dislocation without movement. *Natural Language & Linguistic Theory* 25, 485–534.
- Dik, Simon C.** 1978. *Functional Grammar*. Amsterdam: North-Holland.
- Du Bois, John.** 1980. Beyond definiteness: The trace of identity in discourse. In Wallace Chafe (ed.), *The Pear Stories: Cognitive, Cultural and Linguistic Aspects of Narrative Production*, 230–274. Norwood, NJ: Ablex.
- Etelämäki, Marja.** 2006. *Toiminta ja tarkoite: Tutkimus suomen pronominista tämä* [Activity and referent: A study of the Finnish pronoun *tämä*]. Helsinki: Suomalaisen Kirjallisuuden Seura.
- Etelämäki, Marja.** 2009. The Finnish demonstrative pronouns in light of interaction. *Journal of Pragmatics* 41(1), 25–46.
- Etelämäki, Marja & Minna Jaakola.** 2009. “Tota” ja puhetilanteen todellisuus. [‘Tota’ and the reality of speech situation]. *Virittäjä* 113(2), 188–212.
- Fretheim, Thorstein** 1995. Why Norwegian right-dislocated phrases are not afterthoughts. *Nordic Journal of Linguistics* 18(1), 31–54.
- Geluykens, Ronald.** 1987. Tails (right dislocations) as a repair mechanism in English conversational discourse. In Jan Nuyts & George de Schutter (eds.), *Getting One’s Words into Line: On Word Order and Functional Grammar*, 119–129. Dordrecht: Foris.
- Geluykens, Ronald.** 1992. *From Discourse Process to Grammatical Construction: On Left-dislocation in English*. Amsterdam: John Benjamins.
- Goodwin, Charles.** 1996. Transparent vision. In Elinor Ochs, Emanuel Schegloff & Sandra Thompson (eds.), *Interaction and Grammar* (Studies in Interactional Sociolinguistics 13), 370–404. Cambridge: Cambridge University Press.
- Goodwin, Charles & Marjorie Harness Goodwin.** 1987. Concurrent operations on talk: Notes on the interactive organization of assessment. *IPRA Papers in Pragmatics* 1(1), 1–54.
- Hakulinen, Auli, Maria Vilkuna, Riitta Korhonen, Vesa Koivisto, Tarja Riitta Heinonen & Irja Alho.** 2004. *Iso suomen kieliooppi* [A comprehensive grammar of Finnish]. Helsinki: Finnish Literature Society.
- Halliday, Michael A. K.** 1967. Notes on transitivity and theme in English: Part II. *Journal of Linguistics* 3, 199–244.
- Hanks, William.** 1992. *Referential Practice: Language and Lived Space among the Maya*. Chicago, IL: The University of Chicago Press.
- Helasvuo, Marja-Liisa.** 2001. *Syntax in the Making: The Emergence of Syntactic Units in Finnish Conversation*. Amsterdam: John Benjamins.
- Karhu, Anna.** 1994. Suomen itämurteiden oikealle lohkeavat rakenteet ja niiden lähirakenteet: dislokaatioita, korviketopiiikkeja, diskurssinmerkitsimiä ja pragmaattisia partikkeleita [The right-dislocated constructions and constructions close to them in the Eastern dialects of Finnish: Dislocations, topic substitutes, discourse markers and pragmatic particles]. Licentiate thesis, University of Joensuu.
- Kielitoimiston sanakirja* [A new dictionary of modern Finnish]. 2018. Helsinki: Institute for the Languages of Finland.
- Laitinen, Lea.** 2005. *Hän*, the third speech act pronoun in Finnish. In Laury (ed.), 75–106.
- Lambrecht, Knud.** 1994. *Information Structure and Sentence Form: Topic, Focus and the Mental Representations of Discourse Referents* (Cambridge Studies in Linguistics 71). Cambridge: Cambridge University Press.
- Larjavaara, Matti.** 1990. *Suomen deiksis* [Finnish deixis]. Helsinki: Finnish Literature Society.
- Laury, Ritva.** 1997. *Demonstratives in Interaction: The Emergence of a Definite Article in Finnish* (Studies in Discourse and Grammar 7). Amsterdam: John Benjamins.
- Laury, Ritva.** 2005a. First and only: Single mention pronouns in spoken Finnish. In Laury (ed.), 56–74.
- Laury, Ritva (ed.).** 2005b. *Minimal Reference: The Use of Pronouns in Finnish and Estonian Discourse*. Helsinki: Finnish Literature Society.
- Leech, Geoffrey.** 1983. *Principles of Pragmatics*. London: Longman.
- Maynard, Douglas.** 1980. Placement of topic changes in conversation. *Semiotica* 30(3), 263–290.
- McCarthy, Michael & Ronald Carter.** 1997. Grammar, tails, and affect: Constructing expressive choices in discourse. *Text* 17(3), 405–429.

- Östman, Jan-Ola.** 1995. Recasting the deictic foundation, using physics and Finnish. In Masayoshi Shibatani & Sandra Thompson (eds.), *Essays in Semantics and Pragmatics: In Honor of Charles J. Fillmore*, 247–278. Amsterdam: John Benjamins.
- Pekarek Doehler, Simona, Elwys de Stefani & Anne-Sylvie Horlacher.** 2015. *Time and Emergence in Grammar: Dislocation, Topicalization and Hanging Topic in French Talk-in-interaction* (Studies in Language and Social Interaction 28). Amsterdam: John Benjamins.
- Priiki, Katri.** 2015. *Se oli ihan hullu se hammaslääkäri*. Kaakkois-Satakunnan henkilöviitteiset *se, hän, tää ja toi* eteenpäin lohkeavan konstruktion osina ja ensimmäintoina [Third-person pronouns *se, hän, tää* and *toi* in right-dislocation constructions and as a first mention of a person]. *Puhe ja kieli* 35(2), 47–71.
- Priiki, Katri.** 2017a. *Hän, se, tää vai toi? Vuorovaikutussociolingvistinen tutkimus henkilöviittauksista Kaakkois-Satakunnan nykypuhekielessä* [An interactional sociolinguistic study of third-person references to people in contemporary spoken Finnish in southwestern Satakunta]. Ph.D. thesis, University of Turku.
- Priiki, Katri.** 2017b. The Finnish logophoric pronoun *hän*: A quantitative approach. *Eesti ja soome-ugri keeleteaduse ajakiri. Journal of Estonian and Finno-Ugric Linguistics*, 8(2), 327–349.
- Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech & Jan Svartvik.** 1972. *A Grammar of Contemporary English*. London: Longman.
- Reile, Maria, Piia Taremaa, Tiina Nahkola & Renate Pajusalu.** 2019. Reference in the borderline of space and discourse: A free production experiment in Estonian, Finnish and Russian. *Linguistica Uralica* 55(3), 185–208.
- Schegloff, Emanuel.** 1996. Some practices for referring to persons in talk-in-interaction. In Barbara Fox (ed.), *Studies in Anaphora*, 437–485. Amsterdam: John Benjamins.
- Seppänen, Eeva-Leena.** 1998. *Läsnäolon pronominit: tämä, tuo, se ja hän viittaamassa keskustelun osallistujaan* [Pronouns of presence: *tämä, tuo, se* and *hän* referring to a participant of a conversation]. Helsinki: Finnish Literature Society.
- Siemund, Peter.** 2018. *Speech Acts and Clause Types: English in a Cross-linguistic Context* (Oxford Textbooks in Linguistics). Oxford: Oxford University Press.
- Timmis, Ivor.** 2009. ‘Tails’ of linguistic survival. *Applied Linguistics* 31(3), 325–345.
- Vatänen, Anna.** 2014. *Responding in Overlap: Agency, Epistemicity and Social Action in Conversation*. Ph.D. thesis, University of Helsinki.
- Vilkuna, Maria.** 1989. *Free Word Order in Finnish: Its Syntax and Discourse Functions*. Helsinki: Finnish Literature Society.
- Vilkuna, Maria.** 1992. *Referenssi ja määräisyys suomenkielisten tekstien tulkinnassa* [Reference and definiteness in interpreting Finnish texts]. Helsinki: Finnish Literature Society.
- Ward, Gregory, Betty Birner & Rodney Huddleston.** 2002. Information packaging. In Rodney Huddleston & Geoffrey K. Pullum et al., *The Cambridge Grammar of the English Language*, 1363–1448. Cambridge: Cambridge University Press.
- Winkle, Claudia.** 2015. *Non-canonical Structures, they Use Them Differently: Information Packaging in Spoken Varieties of English*. Freiburg: Albert-Ludwigs-Universität.
- Ziv, Yael & Barbara Grosz.** 1998. Centering, global focus, and right dislocation. In Marilyn A. Walker, Aravind K. Joshi & Ellen F. Prince (eds.), *Centering Theory in Discourse*, 293–307. Oxford: Clarendon Press.

Appendix A. Data sources

- Arksysen: 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. Transcripts available via the Korp corpus search interface of Kielipankki (The Language Bank of Finland) <https://korp.csc.fi/>.
- Satakunta corpus (Sapu): Recordings 154, 155, 157, 159, 161, 163, 164, 165, 167, 187, 188, 190, 202, 209, 210, 211, 222, 223, and 224 from research project Satakuntalaisuus puheessa ‘Satakunta in Speech’. The Syntax Archives at the University of Turku. Department of Finnish and Finno-Ugric Languages, University of Turku.

Appendix B. Transcription symbols

.	falling intonation
,	level intonation
?	rising intonation
↑	step up in pitch
°speak°	quiet talk
£speak£	smiley voice
@speak@	altered voice
<speak>	slowing down
>speak<	speeding up
[beginning of overlap
]	end of overlap
sp-	word cut off
.h	audible inhalation
h	audible exhalation or laughter
speak:	prolonged sound
(.)	micropause (less than 0.2 seconds)
(0.6)	pause length in tenths of a second
(speak)	item in doubt
boldface	a placeholder pronoun in the transcript
<i>italics</i>	a tail NP in the transcript
<u>underline</u>	a constituent in the theme position in the transcript