Pronoun Use in Finnish Reported Speech and Free Indirect Discourse: Effects of Logophoricity

Elsi Kaiser

Abstract Many languages have logophoric pronouns which refer to the person whose speech, thoughts or feelings are being reported, and some languages also have antilogophoric pronouns. This paper investigates (anti)logophoricity in the pronominal system of Finnish, in particular in reported speech and free indirect discourse (FID). I first show that the referential patterns exhibited of two types of third person pronouns in Finnish – the human third-person pronoun hän (he/she) and the non-human third person pronoun se (it), which can also be used for human antecedents in certain contexts – seem to be very different in reported speech vs. FID contexts. However, I argue that the hän/se variation can be derived from a basic generalization – namely that hän refers to SELF (see also Laitinen L, From logophoric pronoun to discourse particle. A case study of Finnish and Saami. In: I Wischer & G Diewald (ed) New reflections on grammaticalization. John Benjamins, Amsterdam/Philadelphia, pp 327–344, 2002) – as long as we take into account (i) the size of the logophoric domain and (ii) the different defaults of standard Finnish and colloquial Finnish. Furthermore, I suggest that we do not need to posit an additional association between se and NON-SELF, because the referential behavior of se can be derived from the size of the logophoric domain and the register defaults. In addition, once we look at how these two pronominal forms interact with the demonstrative pronoun tämä (this) in Finnish, it becomes clear that theories of reference resolution need to consider both logophoricity and salience.

1 Introduction

Some languages have a distinct class of logophoric pronouns that are used to refer to the ‘subject of consciousness’, i.e. the person whose speech, thoughts or feelings are being reported (e.g. Clements 1975; Hагège 1974; Sells 1987; Culy 1994). It has also been suggested that some languages have antilogophoric pronouns which

E. Kaiser (✉)
Department of Linguistics, University of Southern California, Los Angeles, CA, 90007, USA
e-mail: emkaiser@usc.edu

© Springer International Publishing AG 2018
P. Patel-Grosz et al. (eds.), Pronouns in Embedded Contexts at the Syntax-Semantics Interface, Studies in Linguistics and Philosophy 99,
DOI 10.1007/978-3-319-56706-8_3
cannot refer to the subject of consciousness (e.g. Culy 1997). This paper explores the role of logophoricity and antilogophoricity in the pronominal system of Finnish, in particular in contexts involving reported speech and free indirect discourse.

In Finnish, there is more than one pronominal option for referring to human third person antecedents. As illustrated in the examples below from a novel by Anni Polva, both the pronoun *se* ‘it’ and the gender-neutral personal pronoun *hän* ‘s/he’ can be used. In (1a), *hän* is used to refer to the main character, Tiina, and in (1b), *se* is used to refer to the same character. In this example, and elsewhere in the paper, I translate *se* as ‘it’ in English, in order to distinguish it from *hän*.

(1a)  **Tiina** juoksi kotiin niin nopeasti kuin jaloillaan pääsi. **Hän**Tiina hengitti puuskuttaen...

‘**Tiina** ran home as fast as her legs would carry her. **She**Tiina was out of breath . . .’

(Polva 2011: 7)

(1b)  Sekasotkua **se**Tiina joka tapauskesssa oli saanut aikaan, kuten tavallisesti.

‘In any case, **it**Tiina had made a mess of things, as usual.’

(Polva 2011: 29)

This variation raises the basic question of what guides the choice of one form over the other? In this paper, I consider three possible explanations: (i) a register-based account based on the different anaphoric paradigms of standard Finnish and colloquial Finnish/spoken dialects, (ii) a prominence-based account based on the claim that different referential forms refer to antecedents with different levels of prominence in the discourse and (iii) an account related to the notion of logophoricity that builds on the observation that the personal pronoun *hän* ‘s/he’ has a special logophoric use in reported speech in spoken Finnish dialects (Setälä 1883; Kuiri 1984; Laitinen 2002, 2005, inter alia). It will become clear over the course of the discussion that although the first account is not incorrect, it is not sufficient to explain the use of *hän* and *se*. The second account, based on prominence, does not receive any support from the data. The third account is the most promising, and I use it as the foundation for my approach.

After reviewing these three possibilities, I discuss the use of *se/hän* in free indirect discourse and illustrate how (at least at first glance), the referential properties of both *hän* and *se* appear to change depending on whether we are dealing with reported speech or free indirect discourse. However, I claim that we do not need to assume that *hän* and *se* each have two different sets of referential properties. Instead, I show that we can derive the referential behavior of these forms from a single generalization as long as we take into account (i) the fact that reported speech and free indirect discourse are ‘logophoric’ in that they involve reporting a person’s thoughts/speech but differ in the size of their logophoric domains, (ii) the fact that the default pronominal forms are different in standard Finnish and colloquial Finnish and (iii) the fact that FID, by its very nature, mimics properties of spoken language. The relevant basic generalization is that *hän* refers to the logophoric antecedent (see Setälä 1883; Kuiri 1984; Laitinen 2002, inter alia for earlier discussion). Furthermore, I show that we do not need to posit an additional
association between *se* and anti-logophoricity (contra Kaiser 2008), and that the referential behavior of *se* can be derived from the three observations above (see also Hinterwimmer and Bosch 2016 for related data on German d-pronouns, which dislike logophoric referents).

The discussion of *hän* and *se* also brings us to another referential form, the demonstrative *tämä* ‘this’, which can also be used for humans in Finnish. Comparing *tämä* and *se* allows us to investigate the relation between prominence/salience and logophoricity. Many researchers have found that anaphoric forms are sensitive to the prominence of their antecedents. How does this prominence sensitivity interact with logophoricity? As we will see, in contexts with multiple non-logophoric referents, prominence guides the use of *se* and *tämä*. Thus, in order to capture the referential properties of these forms, we need to consider both salience and (non/anti-)logophoricity.

In order to better understand the role of ‘default forms’, I also consider reference to non-human animals, which by default are referred to with *se* ‘it’. What happens when a nonhuman referent is involved in FID and reported speech? In Finnish, the association between logophoricity and *hän* is able to ‘overcome’ the association between *hän* and human referents: *hän* – typically regarded as the human third person pronoun – can be used for animals if the animal is conceptualized as the character whose thoughts are being conveyed, even in contexts where use of *hän* cannot be attributed to personification (see Laitinen 2002).

As a whole, the data presented in this paper highlight the need for models of reference resolution that integrate different kinds of information, including prominence, (non/anti-) logophoricity, and register variation, and are in line with the form-specific, multiple-constraints approach proposed by Kaiser and Trueswell (2008).

### 1.1 Finnish Third-Person Pronominal Paradigm

Before we can start to explain the pronoun variation patterns in ex.(1), let us consider some background information about the pronominal paradigms of Finnish, in particular the split between standard and colloquial Finnish. Standard Finnish is the ‘official’ form of the language and used in formal writing (e.g. newspapers, magazines, non-fiction, textbooks, some fiction) and public/official speech (TV newscasts, speeches etc.). Standard Finnish is also the form on which dictionaries are based. However, in more informal spoken communication and casual writing, people use dialects of colloquial Finnish, which differ from standard Finnish in various aspects of their lexicon, morphology, syntax and phonology/phonetics (e.g. Karlsson 1999). There exist various regional dialects of colloquial Finnish, but as discussed below, the phenomena relevant to us here occur in the majority of dialects and thus (for ease of presentation) I group these together under the label of ‘colloquial Finnish.’ The vast majority of native Finnish speakers can be described
as bidialectal in that they can produce and comprehend both standard Finnish and at least one dialect of colloquial Finnish.

In standard Finnish, third-person human referents are referred to with the gender-neutral personal pronoun hän ‘s/he’ (ex. 1a). Non-human animals and inanimates are referred to with se ‘it’ (ex. 2a, b). Although I translate se as ‘it’, se is often regarded as somewhat of a hybrid that has properties of both anaphoric and demonstrative pronouns (e.g. Larjavaara 1990). In contrast to the proximal demonstrative tämä ‘this’ that expresses proximity to the speaker and the distal demonstrative tuo ‘that’ that expresses distance from the speaker, se has been analyzed as placing the referent in the addressee’s sphere and being unmarked/neutral with respect to the speaker (see Laury 2005). Se can also occur on its own or as a prenominal modifier, in which case its meaning is similar to English ‘the’ or ‘that’, as in ex(2c) (see Laury 1997 on how se is becoming grammaticalized in some dialects as a kind of definite article. Se is also used for discourse deixis (Hakulinen and Karlsson 1989: 316).

(2a) Kissani, nukkuu suurimman osan päivästä. Aamulla se, kuitenkin herää aina samaan aikaan kuin minäkin.
   ‘My cat sleeps most of the day. In the morning, though, it always wakes up at the same time as me.’

(2b) Ostin uuden hienon kissanleluen. Se on täytetty kissanmuntulla.
   ‘I bought a fancy new cat toy. It is filled with catnip.’

(2c) Se hieno lelu oli aika kallis.
   ‘The/those fancy toy was quite expensive.’

Furthermore, even in Standard Finnish se can also be used to refer to humans in certain contexts, in particular in otherwise ‘headless’ relative clauses (ex. 3). Here, se is not used in a typically anaphoric manner, but the fact that it can be the head of a relative clause with a human referent suggests that its features may not be entirely incompatible with human referents.

(3a) Se voittaa, joka ensimmäisenä on purjehtinut 100 meripieninkulmaa.

(adapted from Hakulinen and Karlsson 1988: 314)

It-NOM win-3sing, who-NOM first is sailed 100 nautical-mile-PART
   ‘The first one to sail 100 nautical miles, wins.’

(3b) Pekka on se, jota etsit.

   Pekka-NOM is it-NOM, who-PART look-for-2sing
   ‘Pekka is the one you are looking for.’ (Sulkala and Karjalainen 1992: 120)

In sum, although it might be easy to describe se simply as the inanimate pronoun in Standard Finnish, it is a rather multi-functional hybrid element that has properties of demonstratives, anaphors and determiners, and that can, in some constructions, have human referents.

So far we have been focusing mostly on standard Finnish. The anaphoric paradigms in dialects of colloquial Finnish are quite different. In the majority
of regional dialects, *se* is the default form for human antecedents (as well as animals and inanimates), with the exception of some south-eastern and south-western dialects (e.g., Vilppula 1989). Indeed, Kallio (1978: 65, cited by Suonperä 2012) states that the use of *se* when referring to human referents is so frequent in colloquial language, excepting only the most formal settings, that no specific proof is needed for this observation.

In sum, there is a tension between the pronominal systems of standard Finnish and colloquial Finnish: Whereas *hän* ‘s/he’ is the default pronoun for human antecedents in standard Finnish, *se* ‘it’ is the default pronoun for human antecedents in colloquial dialects. Given that the vast majority of Finnish speakers are bidialectal – i.e. can produce and comprehend both Standard Finnish and at least one dialect of colloquial Finnish – this means that Finnish speakers have, in some sense, two distinct grammatical systems for reference to human antecedents. (In this paper, I focus on the singular forms *hän* ‘s/he’ vs. *se* ‘it’, but it seems that the plural forms *he* (they human) vs. *ne* (they non-human) show the same patterns.) Now, armed with this background, let us return to the pronoun alternation illustrated in ex. (1).

2 A Register Difference?

Having considered the differences between standard Finnish and colloquial Finnish, let us now turn to what might, at first glance, seem like the most straightforward account for the *hän/se* alternation in (1): a register difference. As we saw above, in standard Finnish, the *hän* ‘s/he’ is the default for humans, but in colloquial Finnish, *se* ‘it’ is the default:

(4) Antti tuli eilen kotiin keskiyön jälkeen. Hän/standard se/colloquial nukkui nyt.
    ‘Antti came home yesterday after midnight. He/it is sleeping now.’

Based on these differences, one might be tempted to describe the distribution of *hän* vs. *se* as register-driven, dependent on whether the utterance is in standard or colloquial Finnish. However, it rapidly becomes clear that a purely register-based story is insufficient. As illustrated in (5a,b), from novels by Hannu Raatila and Antti Tuuri, and (5c) from colloquial Finnish, alternating forms (for the same referent) can be used within one register.¹ (*Hiän* in (5c) is a dialectal form of *hän*.) So, though

¹One could try to maintain a register-based account by claiming that these kinds of examples involve register shifts in mid-sentence, such that the utterance starts out entirely in colloquial Finnish (and uses *se*) and then shifts entirely into standard Finnish for the rest of the utterance (and uses *hän*). However, such a ‘full-blown’ register shifting approach does not seem to be appropriate. Intuitively, there is no sense of a full register shift/formality shift here. There are also no morphological or phonological indications of register change. (In Finnish, certain sound combinations change somewhat depending on register, and thus can be used as a tool to detect register shifts.) However, it is worth emphasizing that lack of a ‘full-blown’ register shift does not have to preclude the possibility of referential patterns from one register being borrowed into
the forms differ in their default formality level, the use of hän vs. se cannot be satisfactorily explained by register differences alone.

(5a) **Snell**i ei käsittänyt kuinka tavarat saataisiin pois torilta proomulla. **Se**i kyseli oliko **hän**, ymmärtänyt oikein. (Raittila 2003: 115)

‘**Snell**i didn’t understand how the objects could be transported away from the square by tugboat. **It** asked whether **she** had understood (the plans) correctly.’

(5b) Sanoin, että voisimme vaihtaa paikkoja, mutta sitä **vanha mies** ei halunnut; **sen** mielestä ikkunapaikan saaminen oli kuin arpaajaisvoitto, eikä **hän** halunnut ottaa minulta sitä voittoa pois. . . . **Se** esitteli minulle kameraansa. (Tuuri 1993: 16).

‘I said that we could change seats, but the **old man** didn’t want that. According to it, getting a window seat was like winning the lottery, and **he** didn’t want to take that away from me. . . . **It** showed me its camera.’

(5c) **se** sano jotta kyllä **hän**, suapi tämän paranemmaan. (Kuiri 1984: 120)

‘**It** said that **he** will indeed get this to improve.’

It is important to point out that Finnish novels differ in terms of whether they use se or hän as the default pronoun for human referents. Some of the examples we encounter from novels pattern like colloquial Finnish in the sense of having se as the default form for humans – for example, ex. (5c) is a spoken example from a dialect of Finnish, and shows the same pattern as in ex. (5a) from a novel where se being used in the matrix clause and hän being used in the embedded clause. Thus, in this paper, we are not defining register simply in terms of whether the example comes from spoken Finnish or written Finnish, since written Finnish in novels may be written in a colloquial style (and spoken Finnish could be very formal and employ the standard version of the language). Instead, for all data sources, one should independently assess what the default form is in that particular case.

### 3 A Prominence Difference?

Another factor that might be behind the choice of hän ‘s/he’ vs. se ‘it’ is the salience/prominence of the antecedent. Perhaps one form is used to refer to the most salient referent, and the other is used for less salient referents? For example, given that se has some demonstrative-like qualities whereas hän is more clearly anaphoric, we might expect, based on hierarchies such as Ariel (1990), that hän would prefer more salient/prominent antecedents than se. In order to test this idea, we need some way of measuring salience. Prior research suggests that entities realized in subject another register. We return to this idea in Sect. 6 (see also Maier 2015 for related work). The key point of the Sect. 2 is simply that register alone does not capture the properties of the hän/se variation.
position are more salient than those in object position (e.g. Chafe 1976, Brennan, Friedman and Pollard 1987, inter alia), and so we might expect to see correlations between choice of hän vs. se and the antecedent’s grammatical role. However, this expectation is not supported by the corpus data I have examined. Both hän and se can be used to refer to preceding subjects (e.g. ex.(1a), (5a,b)). Further, as exemplified by (6a) from Raitilä (2003: 216), and (6b) from Raitilä (2003: 44), both hän and se can also be used to refer to non-subjects. Based on the corpus I examined (see Sources section at the end of the paper), there doesn’t seem to be any straightforward link between the grammatical role of the antecedent and choice of hän vs. se.

(6a) Pyysin dosenttia₁ väistämään. Se₁ meni sängyllle makaamaan.  
‘I asked the lecturer₁ to get out of the way. It₁ lay down on the bed.

(6b) Huomautin dosentille₁, että hän₁ oli yhtäkkää tuonut puheeseensa kolmen henkilön nimet aivan kuin ne olisivat yhteisiä tuttujamme.  
‘I told the lecturer₁ that he₁ had suddenly mentioned in his tale the names of three people as if they were our shared acquaintances.’

Additional evidence for the lack of connection between hän vs. se and the salience of the antecedent comes from a translation comparison that I conducted with Raitilä’s novel Canal Grande which has been translated into German by Stefan Moster (2006, BTB Verlag). Out of 12 randomly-chosen occurrences of hän, all 12 are translated into German with a personal pronoun (er ‘he’ or sie ‘she’). Out of 33 randomly-chosen occurrences of human-referring se, 32 are translated into German with a personal pronoun (er/sie, ex(7a-b)) and one is translated with the demonstrative form dieser ‘this one’ (ex.(7c-d)). (Translations into English are by me.)

(7a) German translation
Der Dozent₁ sah mich über seine kleine Brille hinweg wie ein Vorschullehrer. Er₁ würde es schon dazu sagen, wenn er₁ eine Jahresszahl vor Christus meinte. Überhaupt empfahl er₁ mir, mein Gehirn etwas anzustrengen und ihm₁ zu folgen. (Raitilä, German translation, pp. 30–31)
‘The lecturer₁ looked at me over his small glasses like a primary school teacher. He₁ would let me know if he₁ meant a date before Christ (B.C.). He₁ recommended that in general I try to use my brain and follow what he₁ is saying.’

(7b) Finnish original
Dosentti₁ katsoi silmälasiisen yli kuin pikkukoulun opettaja. Hän₁ kyllä sanoo silloin, kun vuosiluku merkitsee aikaa ennen Kristusta. Se₁ käski minua muutenkin käyttämään vähän aivojani ja pysymään mukana. (Raitilä, Finnish original, p. 29)
‘The lecturer₁ looked over his glasses like a primary school teacher. He₁ will specify, when the date is Before Christ (B.C.). It₁ told me to use my brain a little, in general, and to follow along.'
German translation
Plötzlich schnauzte der Bootsführer Heikkilä an. Verlegen hörte dieser sofort auf, über die Geschichte der unsichtbaren Häuser am Kanal zu dozieren. (Raittila, German translation, p. 10)
‘Suddenly the boat captain snapped at Heikkilä. Embarrassed, he immediately stopped lecturing about the history of the invisible houses on the canal.’

Finnish original
Yhtäkkiä venekusi ärähti jotain Heikkilälle. Nolona se lakkasi selostamasta kanavanvarren näkymättömien talojen historiaa. (Raittila, Finnish original, p. 10)
‘Suddenly the boat captain grumbled something to Heikkilä. Embarrassed, it immediately stopped lecturing about the history of the invisible houses on the canal.’

It has been suggested that in German, personal pronouns refer to topics and so-called d-pronouns (der/die) refer to non-topics (e.g. Bosch and Umbach 2006). One might also expect that non-topics could be referred to with the group of longer diese(r/s) demonstratives. However, we find an overwhelming preference for both Finnish forms to be translated into German with personal pronouns (100% of hän and 97% of se). Thus, there is no clear evidence that the choice between hän and se is determined by the salience/topicality of the referent.

4 Use of hän ‘s/he’ in Reported Speech

The preceding sections showed that register differences between standard Finnish and colloquial Finnish do not fully capture the choice of hän vs. se, and that the use of these two forms does not appear to be conditioned by the salience/prominence of the antecedent (at least not if we probe in terms of grammatical role or by comparing pronoun patterns in Finnish and German). In this section we consider a third account, which has been discussed in prior literature on Finnish dialects, and which hinges on the observation that hän has a special use in reported speech in many varieties of colloquial Finnish.

In addition to using se as the default third person pronoun for human referents, many dialects of Finnish use hän in a specific, restricted contexts — namely in reported speech/thought (e.g. Setälä 1883, Kuiri 1984, Ylikahri 1996, Laitinen 2002, 2005, inter alia). Laitinen (2002) calls this a logophoric use, and I will follow her in using this term. A logophoric pronoun is one that refers to the subject of consciousness, the entity “whose speech, thought, feelings or general state of consciousness are reported” (Clements 1975: 141; term coined by Hågège 1974). Adapting a term used by Sells (1987), we can say that a logophoric pronoun refers to SELF. For Finnish, Laitinen (2002) notes that the pronoun hän “appears in reported
speech or thought and is coreferential with the subject of the speech act or mental verb used to introduce it” (Laitinen 2002: 327). For example, in (8a), hän ‘s/he’ is used inside the reported speech context to refer to the matrix subject (realized with se ‘it’) whose speech is being reported. The same pattern can be seen in (8b,c). These examples are all from colloquial Finnish (signaled by morphological patterns and other grammatical cues).

(8a) [Context: talking about good fishing spots]
Kundi₁ luulee omistavansa sen paikan, vaikka mä olin aamulla jo tuntia ennen sitä sillä paikalla. Se₁ sano, että hän₁ on tään paikan alun perin löytänyt. (www.jippii.fi/jsp/forum/thread.jsp?b=kalastus&t=570)
‘The guy₁ thinks he owns the place, although I was already there in the morning an hour before it. It₁ said that he₁ had originally found this place.’

(8b) [Context: talking about good places to use a metal detector to find jewelry] kerroin kouluttajalle₁, että on tosi huonoja rantoja kun on vaan yksi kulta löytynyt ni se₁ sano, että hän₁ tietää yhden hyvän rannan missä käy usein rikkaita
‘I told the trainer₁ that these are really bad beaches since only one gold object has been found so it₁ said that he₁ knows a good beach where rich people often go’

(8c) [Context: waiting to hear back about a possible job as a tractor driver]
kohtra vissiin sen miehen₁ pitää soittaa . . . että koska meidän pitää tavata...
sen tiiän että täänään mutta se₁ sano että hän₁ soittelee läheimpänä viitää
‘the man₁ should probably call soon . . . about when we should meet . . . I know it’s some time today but it₁ said that he₁ will call closer to five o’clock’

4.1 Types of Reported Speech/Thought Configurations

In addition to the straightforward embedding contexts seen in the preceding section, hän also occurs in more complex reported speech/thought contexts. For example, in (9a), hän is in a relative clause embedded inside the subordinate clause under the speech verb in the matrix clause. Thus, the speech verb does not need to be in the immediately higher clause. Relatedly, (9b) shows that the speech verb that embeds the clause with hän does not need to be the highest/matrix clause of the sentence. (These examples are from a novel by Antti Tuuri where se is the default pronoun for humans.) Nevertheless, in these kinds of examples, hän is in the scope of the speech/mental verb in a higher clause and coreferential with the subject of that speech/mental verb.
(9a)  [Context: Nick had too much to drink the night before, and a conversation over breakfast the following day reveals that he does not remember all the events of the preceding evening]

Sei kysyi, olisiko eilen minun mielestäni tapahtunut jotakin, jota hän ei ollut rekisteröinyt muistiinsa.

(Tuuri p. 56)

‘It asked whether in my opinion something had happened yesterday that he had not recorded into his memory’

(9b)  [Context: A group of people, including the narrator, has just met a government minister in Cuba, as part of their attempts to research Ernest Hemingway]

Minäkin kättelin ministerin, joka sanoi minua neidaksi ja pyysi soittamaan hänen, henkilökohtaisesti, jos jotakin vaikeuksia Kuubassa oleskelun aikana ilmaantuisi.

(Tuuri, p. 63)

‘I also shook hands with the minister, who called me miss and told me to call him directly if any difficulties arose during my time in Cuba’

Crucially, hän is not used in all embedded contexts. In a context with an embedded clause that has a third person subject that is not coreferential with the speaker/thinker, hän is not used and instead the default se is employed. This is shown in ex.(9c), where the subject of the matrix sentence is the first-person narrator, but the subject of the embedded clause is Nick, one of the characters in the novel.

(9c)  [Context: The narrator is surprised by how much Nick claims to know about his family.]

Kysyin nyt, paljonko se minusta oikein tiesi ja mistä se tietonsa oli saanut

(Tuuri, p. 24)

‘I asked now, how much it actually knew about me and how it had gotten its information’

Furthermore, se can even be used inside a logophoric domain without being antilogophoric, as noted by Hakulinen et al. (2005). They present the examples in (10b,c) to show that se can be embedded under a verb of saying (or thinking) and can still be coreferential with the subject of saying/thinking.

(10a)  Se katsoi vettä ja siltaa ja sanoi että se yöpyy usein tässä hotellissa työreissullaa. (Hakulinen et al. 2005, p. 1409, example from a novel using ‘se’ as default)

‘It looked at the water and at the bridge and said it often stays in this hotel on business trips.’

(10b)  Se sanoi että se tykkää siitä hirveestä (Hakulinen et al. 2005, p. 1409, from colloquial Finnish)

‘It said that it likes it an awful lot’.
This is important, because it shows that Finnish is not a ‘pure’ logophoric language: According to Culy (1994), many West African languages are pure logophoric languages, meaning that a regular pronoun inside a logophoric domain is antilogophoric, i.e., cannot be coreferential with the person whose thoughts/speech are being reported (Culy 1994: 1080). However, Finnish *se* is not like this, as it can still be coreferential with the speaker/thinker in examples like (10).

### 4.2 Probing Logophoricity with Evaluative Adjectives and Epithets

Given that *se* is not anti-logophoric and can be used under verbs of saying/thinking with a coreferential subject, one might wonder whether, in a context where *se* is the default for human antecedents, there is a difference between sentences with *se* and sentences with *hän* in the embedded clause. In other words, since we have sentences with a *se₁...seᵢ* configuration (ex.10) and sentences with a *seᵢ...hänᵢ* configuration (ex.8-9), is there a difference between them?

In this section, I present data which indicate that, in colloquial Finnish contexts where *se* is the default for humans, use of the more marked form *hän* in contexts of reported speech/thought carries more logophoric meaning than use of default *se*. In particular, it seems that use of *hän* suggests that the speech/thoughts of the logophoric center are more concretely reproduced (closer to the actual verbatim speech/thoughts), whereas *se* seems to allow for a greater level of abstraction. (The issues explored here relate in intriguing ways to the *de se*/*de re* distinction, which is unfortunately beyond the scope of this paper.) This idea can be illustrated with evaluative adjectives and with epithets. Example (11), with the evaluative adjective ‘pretty’, shows that when the embedded sentence does not match the expressive content of what was actually said/thought, *se* seems to be preferred over *hän*. Imagine a context in which the conversation in (11a) has just occurred between Laura and Tiina (example adapted from Potts’ (2003) work on expressive attributive adjectives). Later, Tiina shows the blue vase to her friend Liisa and reports Laura’s comment by saying (11b). In (11b), *se* is preferred, because Laura did not refer to the blue vase as being beautiful. (I use # with *hän* to indicate infelicity, but *hän* is not completely out in this context, it is simply less preferred than *se.*) A variant of the same sentence but without the evaluative adjective is fine with both *se* and *hän* (ex.11c).

(11a) Laura: This blue ceramic vase is really ugly. The orange glass vase is much more stylish. Since I can only fit one of them on my shelf, I plan to throw away the ugly blue vase.

Tiina: But I think the blue vase is beautiful! You shouldn’t throw it away.

Laura: Hey, do you want it? Here, take it, it’s yours.
(11b) Tiina: Se₄ sano että se₄/#hän₄ aikoo heittää tämän kauniin maljakon roskiin!
‘It₄ said that it₄/#she₄ plans to throw away this beautiful vase!’

(11c) Tiina: Se₄ sano että se₄/#hän₄ aikoo heittää tämän maljakon roskiin!
‘It₄ said that it₄/#she₄ plans to throw away this vase!’

A similar effect occurs with epithets. Imagine a context where the conversation in (12a) has occurred between Laura and Tiina, from which we can tell that Tiina considers Mikko an idiot but Laura likes him. Later, Tiina reports part of what Laura said by saying (12b). Here, se seems to be preferred over hän because Laura did not use the epithet that Tiina employs in her report (and would in fact disagree with it). A neutral version is fine with both pronouns (ex.12c).

(12a) Laura: Do you know Mikko Läntinen? I just moved, and now I live right next door to Mikko. We ran into each other and talked for a long while yesterday. We even made plans for a date on Saturday!
Tiina: Oh, the tall guy who works downtown? I think Mikko’s a real idiot.

(12b) Tiina: Se₄ sano että se₄/#hän₄ asuu nyt sen idiootin naapurissa.
‘It₄ said that it₄/#she₄ now lives right next door to that idiot.’

(12c) Tiina: Se₄ sano että se₄/#hän₄ asuu nyt sen Mikon naapurissa.
‘It₄ said that it₄/#she₄ now lives right next door to that Mikko.’

We also see something similar with expressions of locative deixis, as in ex.(12d). Here, the embedded clause uses the expression tääsä (‘here’) which is defined relative to when the reported speech was uttered (i.e. when the person she likes said ‘I am staying here’, note also the use of the present tense) and not relative to the location at which the speaker/writer wrote the sentence. Indeed, it is worth noting that in the preceding clause that does not involve reported speech, the speaker/writer uses siihen ‘there’ (an illative-case-marked form of se), and later on uses tääsä ‘here’ (an inessive-case-marked form of tämä, ‘this’) in the reported speech context. Thus, this example corroborates the idea that use of hän in reported speech contexts is associated with a high level of ‘directness’ in terms of how accurately the reported speech is conveyed.²

²Broadly speaking, the different patterns observed with hän and se in contexts involving evaluative adjectives, epithets and locative deixis relate in interesting ways to the de se/de re distinction. The specifics are unfortunately beyond the scope of the present paper and offer an important avenue for future work.
(12d) [Context: a young woman is talking about how students were asked to line up in school and her crush wanted to stay in his position next to her instead of moving forward in line]

Eli siis poikien pitänen neljä tyttöö eteenään ja sit mun ihastus tuli mun kohdalle ja jai siihen vaikkakin sen ois pitänyt jatkaa vielä matkaa ja se sano, et hän pysyy tässä ja tuuppi muita poikia ohitseen. (http://this-life-is-made-just-for-me.blogspot.com/, blog entry from Dec 2012)

‘So the boys had to move forward by four girls [i.e. stand by the girl four girls ahead of where they were] and then my crush got to where I was standing and stayed there even though it should have continued onwards and it said that he stays here and pushed other boys past (him).’

In sum, in reported speech contexts in colloquial Finnish, (i) hän ‘s/he’ acts as a marker triggering a logophoric interpretation and refers to the matrix subject (subject of the speech act or mental verb), and (ii) se ‘it’ is the unmarked pronoun. It might be best described as nonlogophoric but not antilogophoric, given that it can be used in embedded clauses when coreferential with the matrix subject in reported speech contexts.

It is worth noting that, so far, we have focused on reported speech in colloquial Finnish (including written text written in colloquial style/register), which clearly exhibits the hän/se alternation. What about standard Finnish? Reported speech in ‘pure’ standard Finnish does not show the hän/se alternation: se is not used to refer to humans in reported speech, and hän is used in both the main clause and the embedded clause (unlike colloquial Finnish). This can be explained straightforwardly by the fact that hän is the default form in standard Finnish.

However, although we are making progress towards explaining the choice of hän vs. se, this conclusion regarding reported speech contexts does not explain the hän/se alternation in contexts that have no speech act/mental verb, like the examples above in (1) (see also Saukkonen 1967, Hakulinen 1988). In (1a), hän refers to one of the main characters, Tiina. In (1b), se is used to refer to the same character.

5 Use of se/hän in Free Indirect Discourse

To understand contexts like ex.(1) where hän and se seem to alternate in the absence of speech act or mental verbs, let us consider the notion of free indirect discourse (FID), compared to direct speech and indirect speech/reported speech. In direct speech, e.g. Peter said, ‘I will go home tomorrow’ the words of the speaker are quoted directly and inside the quoted segment, the first-person pronoun refers to the speaker. In indirect speech/reported speech, e.g. Peter said that he would go home tomorrow, the speaker is referred to with a third person pronoun. In free indirect discourse, there is no matrix clause with a verb of speaking/thinking, and instead
the text represents a character’s thoughts directly. A third-person pronoun can be used to refer to the thinker: *Peter was tired of sleeping on Tim’s couch. How could anyone sleep on that old thing, with a mattress as hard as a brick? He would go home tomorrow. No one was going to make him change his mind about that.*

As Saukkonen (1967) notes, in Finnish *hän* is used in FID to refer to the speaker/thinker, who I refer to as the SELF (see Sells 1987). More specifically, my corpus data show the following basic patterns: In free indirect discourse, the person whose thoughts are being represented is referred to with *hän* ‘s/he’. Furthermore, *se* ‘it’, when used, refers to the NON-SELF; i.e. a referent other than the one whose thoughts the free indirect discourse represents (see also Saukkonen 1967; Hakulinen 1988; Kaiser 2008). Intriguing related data is discussed for German d-pronouns (vs. personal pronouns) by Hinterwimmer & Bosch (2016), who note that in German, d-pronouns (*der, die*) cannot refer to the individual whose perspective is being assumed for the sentence that contains the d-pronoun: It seems that German d-pronouns resemble the behavior of Finnish *se* in FID contexts.

The referential patterns of free indirect discourse are exemplified by the following excerpt from a novel (Polva 1989: 60). This is a context where Juha has just seen his girlfriend Tiina in an ice-cream parlor with some other boys, and as he was watching them through the window, Tiina turned around and saw him watching. As indicated by the subscripts, *hän* ‘he’ here is used to refer to Juha, whose thoughts we are hearing, and *se* ‘it’ to Tiina. This passage creates a strong effect of free indirect discourse; the reader sees things from Juha’s perspective. (It is important to note that the examples of free indirect discourse discussed in this paper are from contexts where *hän* is the default form for human antecedents.)

(13a) Juha oli lähtenyt tienemä pitkin harppauksin, mutta kun hän Juha oli varma, että häntä Juha ei nähty enää baarin ikkunasta, hän Juha hiljensi menonsa matelemiseksi. Tiina saisi hänet Juha helposti kiinni, jos se Tiina lähtisi heti liikkeelle, ja tottakai se Tiina lähtisi, siitä hän Juha oli varma.

‘Juha had started walking away with long steps, but when he Juha was sure that he Juha was no longer visible from the bar, he Juha slowed his walking down to a crawl. Tiina could easily catch him Juha, if it Tiina left right away, and of course it Tiina would, of that he Juha was sure.’

The same kind of pattern is exemplified in ex.(13b) from a novel by Joensuu (1983). In this extract, one of the characters, Mikael, is coming downstairs very quietly from his bedroom and listening to see who is at home. Mikael — the SELF, from whose perspective we see and hear things from — is referred to with *hän* ‘he’, and his mother with *se* ‘it’. In ex.(13b), as in ex.(13a), the *hän/se* alternation is no longer constrained by the syntactic frame that was central for the reported speech uses. In FID, both *hän* and *se* can be used in a variety of syntactic contexts.
(13b) Mikael tuli portaat alas niin hiljaa kuin osasi.
   Eteisessä oli hämärää. Häni *MIKÄEL* seiso aloillaan, pidätti henkeä ja
   kuunteli.
   Äiti oli keittiössä. Se *MOTHER* silitti pyykkää. Hajusta sen tiesi – ilmassa oli
   kiva, lämmin haju – ja siitä että äiti hyräili hiljaa. (Joensuu 1983: 31)
   ‘Mikael came downstairs as quietly as possible.
   The foyer was dark. He *MIKÄEL* stood still, held his breath and listened.
   Mother was in the kitchen. It *MOTHER* was ironing. One could tell from the
   smell – there was a nice, warm smell in the air – and from mother’s quiet
   humming.’

If we compare ex. (13a) and (13b) to the examples of reported speech in colloquial
Finnish that we have been focusing on so far, we observe broader syntactic
environments in which the *hän/se* alternation is possible. Unlike the reported
speech examples in the preceding sections where *se* was in a matrix clause with
a speech/thought verb and *hän* was in a subordinate clause, we now observe both
*se* and *hän* in matrix clauses, as can be seen in (13a) and (13b). In *FID*, the
*hän/se* alternation is *not* restricted to explicit reported/speech thought configurations
involving syntactic embedding.

At this point in the paper we can also revisit examples (1a) and (1b), repeated
below. Ex.(1a) comes from the very start of the novel and does not involve *FID*.
Hän is used for reference back to Tiina since it is the default form in standard
Finnish. However, once we take a closer look at the context of example (1b) – added
below in the longer version (1b’) – it becomes clear that this example contains free
indirect discourse from the perspective of Tiina’s mother. In this example, following
an exchange where Tiina and her brother have been telling their parents about what
happened at school, the reader hears the mother’s thoughts about her daughter. The
sense of Tiina being the NON-SELF is especially strong in the second and third
sentences of ex.(1b’) below, where the form *se* ‘it’ is used to refer to Tiina.

(1a) Tiina juoksi kotiin niin nopeasti kuin jaloillaan pääsi. Hän *TIINA*
   hengitti
   puuskuttaen . . .
   (Polva 2011: 7)
   ‘Tiina ran home as fast as her legs would carry her. She *TIINA* was out of
   breath . . .’

(1b’) Äiti ei tiennyt mitä sanoa ja mitä oikein ajatella; olisko Tiina tehnyt jotain
   rangaistavaa, vai olisko hän *TIINA* syytön. Sekasotkua se *TIINA* joka
   tapauksessa oli saanut aikanan, kuten tavallisesti. Siellä missä Tiina oli,
   siellä tapahtui aina jotakin, vaikkei se *TIINA* olisi tehnyt muuta kuin seisonut
   hiljaa paikallaan. (Polva 2011: 29)
   ‘Mother didn’t know what to say or quite what to think, had Tiina done
   something wrong, or was she *TIINA* innocent. In any case, it *TIINA* had made a
   mess of things, as usual. Wherever Tiina was, something was always
   happening there, even if it *TIINA* wasn’t doing anything more than standing
   still.’
However, use of *se* in free indirect discourse for the NON-SELF referent is not obligatory; *hän* can also be used (see Rouhiainen 2000). In (13c), which comes from a longer extract of the Finnish translation of “Women in Love” D.H. Lawrence that is discussed by Rouhiainen (2000: 118), we hear the thoughts of one of the protagonists, Gudrun, about her lover Gerald, and *hän* refers to the NON-SELF referent Gerald. (The larger context of this extract makes it clear that we are heading Gudrun’s thoughts, that she is the SELF.) Perhaps relatedly, in the longer version of ex. (1b’)) given above, *hän* is also used to refer to Tiina, although it is not entirely clear whether this use occurs inside FID or not.

(13c) **Geraldin** pitäisi päästää sellaiseen asemaan, missä **hän** *(Gerald)*
thadonvoimallaan ja ylivertaisella käytännöllisellä älyllään voisi ratkaista
nykyajan teollisuuden pulmat. (Lawrence 1980: 541)
‘**Gerald** should achieve a position where **he** *(Gerald)*, with his force of will
and supreme practice intelligence, would be able to solve the problems of
modern industry.’

These observations are in line with earlier corpus work by Rivinoja (2006), who
found that out of 29 references to SELF in Finnish FID (in novels originally written
in Finnish), the human pronoun *hän* was used 83% of the time and proper names
were used 17% of the time. She did not find any cases of *se* being used to refer to
SELF in FID. Rivinoja also analyzed 46 occurrences of reference to NON-SELF
in Finnish FID and found that 11% are accomplished with *hän*, 35% with a proper
name, 26% with *se*, 24% with nouns of various types and 4% with other kinds of
expressions.

In sum, we see that inside FID contexts, *hän* can be used to refer to SELF
(ex.13a,b) or to NON-SELF (13c) and *se* is used to refer to NON-SELF (13a) but
not to SELF. The seemingly ‘mysterious’ alternation in examples (1a) and (1b) at
the start of the paper can now be attributed to *se* being used for the NON-SELF.

As mentioned earlier, the examples of free indirect discourse discussed in this
paper are from contexts where *hän* is the default form for human antecedents –
i.e., from standard Finnish. When we consider the question of what happens with
free indirect discourse in colloquial Finnish, the picture becomes more complex.
As will become clear in the course of this paper, my approach predicts that in
colloquial Finnish, free indirect discourse will resemble reported speech, in that
inside logophoric domains, *hän* will only be used for reference to SELF. Vilppula
(1989)’s discussion of dialectal corpus data seems compatible with this. However,
further work is needed to test the validity of my prediction in detail. It is also worth
noting that FID is presumably much less frequent in colloquial registers than in
formal literature – thus, in this paper, our discussion of FID is limited to standard
Finnish only.
5.1 Two Forms with Different Behaviors in Different Contexts

The data we have seen so far suggests that when hän and se are used inside reported speech contexts (Sect. 4) in colloquial Finnish, (i) hän is logophoric and (ii) se is unmarked/nonlogophoric in that it can be used to refer the person whose thoughts/speech are being reported, as well as other referents. In contrast, in free indirect discourse (Sect. 5) in standard Finnish, (i) hän is unmarked/nonlogophoric in that it can be used for SELF but also for other referents, but (ii) se is antilogophoric in that it cannot be used to refer to SELF and picks out some other referent. This is summarized in Fig. 1.

To better understand the different uses of hän ‘s/he’ and se ‘it’, let us consider an example where the interpretation of hän depends on whether the sentence is interpreted as reported speech or free indirect discourse. Consider ex.(14a). This could be reported speech: Imagine that Tiina’s mother is talking about Tiina’s travels, including one occasion where Tiina hopped on a train and thought that she hadn’t paid for her ticket (when in fact a friend had paid for it). Here, hän in the embedded clause is coreferential with the matrix subject.

(14a) Se luuli, että hän ei ollut maksanut lippuaan.
   ‘It thought that s/he hadn’t paid for its/her/his ticket.’

(14b) Tiina’s mother says:

   Setiina luuli, että hänTiina ei ollut maksanut lippuaan.
   ‘ItTiina thought that sheTiina hadn’t paid for her ticket.’

Now let us imagine a context where (14a) is inside a stretch of free indirect discourse, as in (14c). Imagine this context: Tiina pays for her train ticket, but sees the train starting to pull out of the station and forgets $10 worth of change on the ticket counter when she starts running towards the train. The person selling tickets calls out after her. Tiina hears someone shouting “Stop! The money!” but she is already on the train when she realizes that it was the ticket seller trying to get her attention. Just then, she sees a policeman hop into the train.
(14c) Tiina watched incredulously as the angry-looking policeman hopped onto
the train and walked directly into the compartment where she was sitting.
Why did this stuff always happen to her?
.... Se POLICEMAN luuli, että hän TIINA ei ollut maksanut lippuaan.
.... ‘It thought that she TIINA hadn’t paid for her ticket.’

Here, we are hearing Tiina’s thoughts, and the pronoun hän in the embedded
sentence refers to her. Se refers to the NON-SELF, i.e. the policeman. Thus, if (14a)
is an example of reported speech (ex.14b), se and hän are coreferential, but if we
interpret the same string of words as being free indirect discourse (ex.14c), then hän
and se are disjoint.

In fact, this ‘minimal pair’ simply highlights a pattern we already saw earlier: We
already saw examples of se and hän coreferring in reported speech contexts in Sect.
4 (e.g. ex.(8a–c)). We already saw disjoint reference of hän vs. se in FID contexts
in examples (13a,b). (When considering these examples, it is important to keep in
mind that in FID contexts, use of hän to refer to SELF and se to refer to NON-SELF
also occurs in contexts (such as (13a,b)) that do not involve syntactic embedding
under verbs of speaking/thinking.) In the current section, ex.(14c) uses a case of
reported thought – embedded inside FID – to create a minimal pair with ex.(14b),
to highlight the difference in how the referential forms are interpreted. But the cases
of FID that we are interested in are not confined to cases of syntactically embedded
reported speech/thought inside FID, as we already saw in ex.(13a,b).

The differences between (14b) and (14c) could be taken to imply a conclusion
where the two forms hän ‘s/he’ and se ‘it’ have different referential properties when
used in FID and when used in reported speech/thought contexts without FID, as
shown in Fig. 1—but is this a desirable conclusion? In Sect. 6, I propose a more
unified approach.

6 Unifying the Referential Behavior of hän and se

In this section, I argue that we can reconcile the seemingly disparate patterns in Fig.
1 by taking into consideration (i) the size of the logophoric domain, (ii) the default
form in the register, and (iii) the fact that FID, by its very nature, mimics properties
of spoken language.

First, let us consider the size of the logophoric domain. If we compare reported
speech (14b) and free indirect discourse (14c), we can see that the logophoric
domain (the part that represents the thoughts/speech of a particular character,
marked with [ . . . ] below) is larger in free indirect discourse (15b) than reported
speech (15a):
(15a) Tiina’s mother: “SeTIINA thought that [hänTIINA hadn’t paid for her ticket].”
    Tiina thinks: “ITIINA haven’t paid for my ticket.”

(15b) Narrator: Tiina: [SePOLICEMAN thought that hänTIINA hadn’t paid for her
ticket].
    Tiina thinks: “HePOLICEMAN thinks ITIINA haven’t paid for my ticket.”

In particular, in reported speech, the matrix subject (realized with se in ex.14b/15a)
is not inside the logophoric domain, but the embedded subject (realized with hän) is
inside the logophoric domain. However, in free indirect discourse (14c/15b), both
the matrix subject (realized with se) and the embedded subject (realized with hän)
are inside the domain. Before saying more about this difference, let us also think
back to the register differences between colloquial and standard Finnish: In standard
Finnish hän is the default third person pronoun, and se is normally only used for
nonhuman referents. In colloquial Finnish, se can be used for human referents,
and in fact is the default third person pronoun. Recall also that our discussion of
FID focuses on standard Finnish, and our discussion of reported speech focuses on
colloquial Finnish.

If we combine these observations about (i) the size of the logophoric domain and
(ii) register defaults, we can represent the referential properties illustrated in Fig. 1
in a more unified way, as shown in Fig. 2.

This approach lets us represent the referential properties that hän and se display
in reported speech in colloquial Finnish and free indirect discourse in standard
Finnish, while requiring only one statement about logophoricity, namely that hän
is associated with reference to SELF. As I will show below, we do not need to
posit an association between se and NON-SELF, as this can be derived from
independent properties of FID. (In earlier work, Kaiser 2008, I proposed a more
complex approach involving two associations, one linking hän to SELF and the
other linking se to NON-SELF. The current proposal does not require the second
association).

Let us now consider how this approach can capture examples like (14) and (15).
In a context where the default pronoun for human antecedents is se (i.e., colloquial
Finnish, including fiction written in the colloquial register), use of hän inside a

Fig. 2  Referential properties
of hän ‘he/she’ and se ‘it’

| (a) Register defaults for reference to humans |
| Standard Finnish: hän |
| Colloquial Finnish: se |
| (b) hän => logophoric/SELF |
logophoric domain is associated with reference to SELF, as in the reported speech example in (15a). The default form *se* can also be used in such contexts (precisely because it is the default form in the register), as shown by examples like (9c,10), but does not appear to carry the same perspectivizing effect, as we saw in ex.(11–12).

These patterns are illustrated in part (a) of Fig. 3, where the (gray) dotted lines show the default mappings that arise because *se* is the default in the colloquial register, and the solid line represents the link between *hän* and SELF. Due to the register default, the result is that *hän* is only used when referring to SELF.

In contrast, in a context where the default pronoun for human antecedents is *hän* (i.e. standard Finnish), *hän* can be used to refer to SELF (in accordance with Fig. 2 part (b)), as we saw in FID examples like (13a). However, *hän* can also be used for reference to NON-SELF, as in examples like (13c), because it is the default form in the register. This is illustrated in part (b) of Fig. 3, with dark dotted lines. I have also included the solid line that reflects the mapping between *hän* and SELF, but since *hän* is the default in this register anyway, there are no detectable effects of the link between *hän* and SELF in Standard Finnish. (Thus, I assume the mapping between *hän* and SELF exists in both registers, simply for reasons of parallelism/simplicity, but the effects are only detectable in the colloquial register).

This brings us to the question of how to capture the fact that in FID examples like (13a), *se* is associated with reference to the NON-SELF. Nothing in Fig. 2 directly links *se* to NON-SELF, so how can we explain this? In earlier work, I argued for an explicit association between *se* and NON-SELF (Kaiser 2008), in addition to the association between *hän* and SELF. However, in this paper I claim that such an association is not necessary: The association between *se* and NON-SELF in FID contexts can be derived by combining the well-known fact that FID mimics spoken language (e.g. Tiittula and Nuoliäärvi 2013 and many others, see also footnote 3), with the fact that in colloquial Finnish, *se* is the default form. This is shown in part (b) of Fig. 3 with the gray dotted line. Crucially, we do not need to posit a special link between *se* and NON-SELF — that link ‘comes for free’ from the default pattern of colloquial Finnish. More concretely, consider ex.(14c) and (15b), repeated below:
(14c) Tiina watched incredulously as the angry-looking policeman hopped onto the train and walked directly into the compartment where she was sitting. Why did this stuff always happen to her?
   
   . . . Sepoliceman luuli, että hänTiina ei ollut maksanut lippuaan.
   . . . ‘ItPoliceman thought that sheTiina hadn’t paid for her ticket.’

(15b) (i) Narrator: Tiina: [Sepoliceman thought that hänTiina hadn’t paid for her ticket].

   (ii) Tiina thinks: “HePoliceman thinks ITiina haven’t paid for my ticket.”

Let’s start with line (ii) of example (15b). Tiina’s original though is that “He thinks I haven’t paid for my ticket.” When this is realized as FID, the SELF-referring “I” is realized as hän. The NON-Self-referring pronoun ‘he’ is realized as se, simply because FID mimics properties of colloquial language and se is the default form in colloquial Finnish. Thus, we end up with a surface form where hän refers to SELF and se to NON-SELF, as shown in (14c) and (15b, i). In sum, I build on the observation that FID mimics properties of spoken language and use that to derive the pattern that in FID contexts, se refers to NON-SELF.\(^3\)

7 Adding a Third Form to the Mix: Demonstrative tämä ‘This’

In addition to hän ‘s/he’ and se ‘it’, human antecedents in Finnish can also be referred to with the proximal demonstrative tämä ‘this’. In this section I show that se and tämä can both be used for reference to NON-SELF, but differ in the salience of their antecedents: When there are multiple NON-SELF referents present, se is used for the most salient one and tämä for less salient ones. Existing research characterizes tämä (in its human anaphoric use) as referring to background characters/nonsalient referents, which are often objects, obliques, etc. (e.g. Varteva 1998; Halmari 1994; Kaiser 2003; Kaiser and Trueswell 2008), as in (16a). Tämä is also used to refer to humans in colloquial Finnish (e.g. Etelämäki 2005), but in this section we focus on standard Finnish.

---

\(^3\)One possible avenue for formalizing this intuition about colloquial Finnish patterns persisting in FID could be Maier’s unquotation analysis (e.g. Maier 2015): If we regard FID as essentially a type of direct speech, with SELF-referring pronouns having been ‘unquoted’, then NON-SELF-referring pronouns could potentially maintain their ‘colloquial properties’ simply by virtue of their origins in direct speech which would be in colloquial Finnish. Under this approach, NON-SELF pronouns would not be unquoted.
(16a) FIA julkaisi keskiviikkona tiedotteen, jossa se kummasteli Ecclestonen lausuntoa. Tämä väitti taistellessa jo vuosia F1-sääntöjen tiukkuutta vastaan. (from the newspaper Helsingin Sanomat 21.20.1999) ‘FIA published on Wednesday an announcement in which it expressed surprise at Eccleston’s statement. This claimed to have been fighting for years against the strictness of the F-1 regulations.’

These kinds of examples do not have any logophoric flavor, and a salience-based account works well. However, a salience-based story is not sufficient for all occurrences of tämä. In certain contexts, tämä is used to refer to a preceding subject or an otherwise salient referent. More specifically, this kind of use is possible in antilogophoric, FID-type contexts (see Varteva 1998). E.g., in (16b), the woman referred to with tämä is described from the miller’s perspective; the miller is SELF. Similarly, in (16c), Tina is described from Antti’s perspective; Antti is SELF.

(16b) [Context: The miller hears someone call his name and turns to look:] Mylläri käänätyi äänen suuntaan. Kaunis nainen seisoi sillalla. Tämä oli riisunut huivinsa ja heilutti sitä kiehtovasti.
‘The miller turned towards the direction of the sound. A beautiful woman stood on the bridge. This had taken off her scarf and waved it in a captivating fashion.’ (Paasilinna 1998: 19)

(16c) [Context: Antti and Tina are sitting in a restaurant, having dinner. But Antti can’t relax; he feels there is something wrong with Tina.]
Tinan pirteydessä oli jotain pakotettua. Tämä yritti peittää jotakin hurmaavuudellaan, mutta silmissä oli oudon surumielinen kätse.
‘There was something forced about Tina’s cheerfulness. This was trying to hide something by being so charming, but her eyes looked strangely melancholy.’ (Remes 2001: 21)

7.1 What Is the Relation Between tämä ‘This’ and se ‘It’?

The use of tämä for NON-SELF referents in standard Finnish perspectiveizing contexts, as illustrated above, brings up the question of how tämä and se relate to each other. Are they simply two functionally equivalent ways of referring to NON-SELF referents? If we follow the approach that ranks referential expressions on a salience/accessibility scale (e.g. Ariel 1990, Gundel, Hedberg and Zacharski 1993), we predict that se ‘it’ will refer to more salient entities than tämä ‘this’, since pronouns are predicted to refer to more salient entities than (anaphoric) demonstratives.

To test this, consider the examples below, which have two NON-SELF referents. These are FID-type contexts, in that we are presented with the thoughts of a character in a narrative, in this case a woman called Tiina. Imagine a situation
like the train-ticket scenario in (14), except that now, instead of seeing an angry policeman hop into the train, Tiina sees an old lady on the train eyeing her curiously, and then the old woman gets up and goes to speak to a policeman who is already sitting on the train. Now, consider (17a). Here, in the free indirect discourse, Tiina is the SELF, and there are two NON-SELF referents: the old lady and the policeman. Assuming that subjects are more salient than objects/obliques (e.g. Chafe 1976, Brennan, Friedman and Pollard 1987), in (17a), the policeman (subject) is, crucially, more salient than the old lady (object possessor). Let us now consider two possible continuation sentences, shown in (17b) and (17c). The subsequent sentence (17b) contains both se and tämä, and verb semantics make it clear that tämä (in subject position) refers to the old lady, se (in oblique object position) to the policeman, and hän to Tiina. Ex. (17c) has the same verbs but now se is in subject position and tämä is the oblique argument:

(17a)  Tiina katseli ihmeissään, kun poliisi kuunteli vanhan rouvan kiihkeää selitystä.
       ‘Tiina looked on, surprised, as the policeman listened to the old lady’s impassioned explanation.’

(17b)  Tämä ilmeisesti selitti sille, että hän matkusti pummilla.
       ‘This was apparently explaining to it that she(Tiina) was traveling without paying.’

(17c)  # Se ilmeisesti selitti tälle, että hän matkusti pummilla.
       ‘It was apparently explaining to this that she(Tiina) was traveling without paying.’

Crucially, while (17b) is judged to be fine, (17c) is infelicitous. In other words, the referential mapping that is felicitous is the one where tämä refers to the old lady (the genitive of the object in the preceding sentence), and se refers to the policeman (the preceding subject). This suggests that se is used to refer to more salient NON-SELF referents than tämä.

When the grammatical roles of the policeman and the woman are reversed, as in (18a), so that the policeman is less salient, the interpretation of tämä and se is also correspondingly reversed. A subsequent sentence where tämä refers to the old lady, se to the policeman, and hän to Tiina, which was judged to be felicitous in (17b), is now infelicitous (18b). In contrast, a sentence where tämä refers to the policeman, se to the old lady, and hän to Tiina, is now felicitous (18c). So, the referential mapping judged to be felicitous is one where tämä refers to the policeman (the object in the preceding sentence), and se refers to the old lady (the subject of the preceding sentence). This supports the observation that in contexts with multiple NON-SELF referents, the referential labor between tämä and se is divided based on salience, with se referring to more salient entities than tämä.

(18a)  Tiina katseli ihmeissään, kun vanha rouva selitti kiihkeästi poliisille.
       ‘Tiina looked on, surprised, as the old lady explained something excitedly to the policeman.’
(18b)  # Tämä ilmeisesti selitti sille, että hän matkusti pummilla.
‘This was apparently explaining to it that she(Tiina) was traveling without paying.’

(18c)  Se ilmeisesti selitti tälle, että hän matkusti pummilla.
‘It was apparently explaining to this that she(Tiina) was traveling without paying.’

These data show that in logophoric contexts—just like in nonlogophoric contexts—referent prominence has an effect on how the NON-SELF referents are referred to. The prominence difference (se > tämä) is revealed in the presence of multiple NON-SELF referents. These patterns are compatible with the approach outlined in Figs. 2 and 3, because tämä is not the default/unmarked third person anaphoric form in either standard or colloquial Finnish. Thus, its use would not ‘override’ the register-related defaults or the association that hän has with logophoricity.

8  Non-human Animates: When the Default Pronoun Is Different Due to [-Human] Feature

So far we have focused on situations where the pronouns hän ‘s/he’ and se ‘it’ refer to humans in reported speech contexts and FID contexts. However, there are also situations where the relevant referent is a non-human animal, for example in children’s stories. These contexts differ from the human contexts in one critical and very relevant dimension: In these contexts, se is the default pronoun for the referent in both standard Finnish and colloquial Finnish, simply because se is the basic non-human pronoun, like English ‘it’ – i.e., se is the default not due to register-related factors but due to a featural property of the referent. By looking at cases where the identity of the default form is due to something other than register, we can get a better understanding of how the general notion of ‘default form’ contributes to the interpretation of hän and se.

It is worth noting right away that, as in English, the traditional ‘se = non-humanhän = human’ mapping can be violated when animals are personified. This is can be easily observed in fiction written in standard Finnish. For example, in the Finnish translation of Paddington Bear by Michael Bond (2015), the pronoun se is initially used when Paddington is first found by Mr. and Mrs. Brown (ex.19a), but as soon as they give him a name, the Finnish translation switches to the pronoun hän (ex.19b). After this point, hän becomes the default pronoun for Paddington.
(19a) [Context: Mrs. Brown takes her first close look at the still-unnamed bear]
Hän tirkisteli karhua tarkemmin. Se ei ollut mikään tavallinen karhu. Se oli ruskea – varsin likaisenruskea – ja sillä oli hyvin merkillinen, leveälirinen hattu juuri niin kuin herra Brown oli sanonut. (Bond, 4/67)
‘She observed the bear more closely. It was not a regular bear. It was brown – rather dirty brown – and it had a very odd, broad-rimmed hat just like Mr. Brown had said.’

(19b) [Context: Right after Mr. and Mrs. Brown have decided to name the bear Paddington, after the station where he was found]
Paddington nuoli huuliaan. – Minulla on hirveä jano, hän sanoi. (Bond, 7/67)
‘Paddington licked his lips. ‘I’m very thirsty,’” he said.’
(Bond 1958/2003, English original, ch. 1)

However, the kinds of contexts we are interested in are stories written in Standard Finnish where humans are by default referred to with hän ‘s/he’ and, crucially, animals are by default referred to with se ‘it’. We want to see what happens in these contexts if and when a nonhuman referent is involved in FID and reported speech. Will we see patterns similar to FID and reported speech in colloquial Finnish, with se being the default and hän being used in logophoric contexts? Or will the non-humanness of the referent block use of the human pronoun hän?

### 8.1 Use of hän with Non-human Animates

In reported speech/thought contexts, as Laitinen (2002) notes, hän can be used for animals, (ex.20). Laitinen emphasizes that this is “not a case of secondary personification” and states that the “referent of the logophoric pronoun (…) hän can be any being whose behavior the speaker is able to understand” (Laitinen 2002: 333). The observation that hän can be used to refer to SELF even when SELF is non-human challenges the view that hän is associated with [+human] referents (or humanized/personified referents).

(20) Mut koera jos ottaa ni se tietää että mihinkä hän viep (example from Laitinen 2002; colloquial Finnish)
‘But if the dog takes (something), it knows where s/he takes (it)’

Similar to what we see in reported speech/thought contexts, corpus data show that hän can be used in FID contexts to refer to the logophoric SELF, even if it is an animal. The examples I discuss here are from a young adult novel called “Bernie ja Tiina” (Kukkanen, 2014). The novel is written in standard Finnish, and the default pronoun for humans is hän and the default for animals is se. The story is about a dog, Bernie, that is sent down from “dog heaven” to help a young girl convince her parents that she should be allowed to get a dog. Crucially for our purposes, the
default pronoun for animals in this novel is _se_. Ex.(21a) shows that even in contexts where Bernie is personified and engaging in mental activities (e.g. thinking and planning an upcoming speech), the nonhuman pronoun _se_ is used:

(21a)  **Bernie**₁ meni suihkuhuoneeseen ja väänsi veden täysillä valumaan.
Mielessän _se_₁ jo suunnitteli puhettaan koko koirayleisön edessä. _Se_₁ mietti myös millainen tarinan loppu voisi olla . . .  
(Kukkanen, p. 50)

‘**Bernie**₁ went into the shower room and turned the water on full. Mentally, _it_₁ was already planning its speech in from of the whole dog community.
**It**₁ also wondered about how everything would turn out . . . ’

Although _se_ is the default for referring to Bernie (and other dogs), there are numerous examples of _hän_ being used for Bernie inside FID contexts, as in (21b,c) below. However, the default _se_ is also used in FID, as in ex.(21d):

(21b)  [Context: Tiina’s mother had commented on the odd size relation between Bernie, who is very big, and the size of his dog food bag, which is rather small]

Taas **Bernietå», ihmettyi. Mitä suhdetta siihen tarvittiin? Ei muuta kuin ruoka kuppiin, niin kyllä _hän_₁ sille suhteita osoittaisi. Suorinta tieta vatsaan ja sillä hyvä  
(Kukkanen, p. 23)

‘**Bernie**₁ was confused again. What kind of relation did that need? Just put the food in the cup, and _he_₁ would show it the right kind of relation. Straight to the stomach and that’s it.’

(21c)  [Context: Bernie has been locked into the kitchen to sleep at night]

Huokaisten **Bernie**, istahti miettimään. Ei ollut ollenkaan mukavaa nukkua yksin keittiössä. Mikä kyökkipiika _hän_₁ muka oli?! Iso ja komea berninpaimenkoira poika . . .  
(Kukkanen, p. 35)

‘With a sigh, **Bernie**₁ sat down to think. It was no fun to sleep alone in the kitchen. What kind of scullery maid was _he_₁? A big and handsome Bernese mountain dog . . . ’

(21d)  [Context: Tiina’s father and brother find an announcement for a dog that went missing in Lahti, and think that maybe Bernie is that missing dog]

**Berniekin**₁ ihmettel. Miten ihmeessä _se_₁ olisi voinut kadota Lahdessa perjantaina, kun _se_₁ oli poistunut vasta sunnuntaina Koirien Taivaasta?  
(Kukkanen, p. 40)

‘**Bernie**₁ was surprised too. How on earth could _it_₁ have disappeared in Lahti on Friday, when _it_₁ had only left the Dogs’ Heaven on Sunday?’

In sum, these FID examples are in line with Laitinen’s (2002) observations from colloquial Finnish, and show that _hän_ can be used logophorically for non-human referents also in standard Finnish.⁴ Nevertheless, the fact that _se_ is also used in

---

⁴In addition to being used to refer to SELF in FID and reported thought contexts, it seems that _hän_ can also be used for non-human animals in other contexts, at least by some authors. E.g., a
Fig. 4 Interplay of the logophoricity of \( h\ddot{a}n \) and the use of \( se \) for non-human referents

FID contexts (ex.21d) shows that it is still the default form for non-human animals. So, we have seen that with non-human referents, \( se \) – which is the default – can be used in non-logophoric as well as in logophoric contexts, but the non-default human pronoun \( h\ddot{a}n \) is only used when referring to SELF, in FID or reported speech contexts. These patterns are illustrated in Fig. 4: The dotted lines show that the default (not due to register, but features of the referent) here is \( se \), since we are talking about reference to animals. As before, the solid line illustrates the association between \( h\ddot{a}n \) and SELF, in the same way as in the preceding diagrams. Crucially, the SELF usage of \( h\ddot{a}n \) cannot be attributed to personification, since contexts that involve personification without FID or that are not in the scope of reported speech/thought do not involve use of \( h\ddot{a}n \) (ex.21a).

Put together, the evidence from animal referents provides strong evidence that \( h\ddot{a}n \) is associated with a logophoric/SELF interpretation, as posited in Figs. 2 and 3. One might thus speculate that perhaps \( h\ddot{a}n \) is not associated with [+human] at all but only with [+SELF], as this would explain why it can be used to refer to human and non-human SELF referents. However, this cannot be the case, since other evidence shows that \( h\ddot{a}n \) is indeed the default form for human referents in standard Finnish even when they are NON-SELF (ex.13c). Instead, it seems that \( h\ddot{a}n \) is associated with both a [+human] feature and a SELF feature, and that the SELF feature is higher ranked/more influential than the [+human] feature: A [-human] SELF referent can be referred to with \( h\ddot{a}n \) thanks to its SELF status.

9 Conclusions

This paper explores the use of different referential forms in Finnish, where humans can be referred to with three different anaphoric forms, \( h\ddot{a}n \) ‘she/he’, \( se \) ‘it’ and \( t\ddot{a}m\ddot{a} \) ‘this’. We took as our starting point the question of what guides the use of \( h\ddot{a}n \) vs. \( se \). Even though it may at first glance look like \( h\ddot{a}n \) and \( se \) change their referential properties depending on whether we are dealing with reported speech or free indirect speech, I argue that the \( h\ddot{a}n/se \) variation can be derived from a basic

---

children’s book by Elina Karjalainen (“Uppo-Nalle ja Setä Tontoni”) mostly uses \( se \) when talking about the main characters – teddy bears and other animals – and uses \( h\ddot{a}n \) only very rarely. Based on my analyses so far, Karjalainen’s uses of \( h\ddot{a}n \) for animals, however, do not refer to SELF, but are nevertheless related to perspective-taking. This merits further research.
generalization, as long as we take into account (i) the size of the logophoric domain and (ii) the different defaults of standard Finnish and colloquial Finnish. The basic generalization is that hän refers to a SELF antecedent (see also Setälä 1883; Kuiri 1984; Laitinen 2002, 2005, inter alia). Furthermore, I show that we do not need to posit an additional association between se and NON-SELF (contra Kaiser 2008), because the referential behavior of se can be derived from the size of the logophoric domain and the register defaults, as long as we keep in mind the basic observation that free indirect discourse mimics properties of spoken language (i.e., the colloquial register).

I also investigate the relation of se to another NON-SELF form, the demonstrative tämä ‘this’, which can also be used anaphorically for human antecedents in Finnish. I conclude that in the presence of multiple NON-SELF referents, the division of labor between tämä and se depends on the prominence of the antecedent, with tämä being used for less prominent antecedents. Thus, to capture the referential properties of the three forms investigated in this paper, we need to consider (i) logophoricity (whether the antecedent is SELF or not), (ii) what the register defaults are, and (iii) how prominent/salient the antecedent is. These findings are in line with the form-specific, multiple-constraints approach proposed by Kaiser and Trueswell (2008), according to which different referring expressions can differ in how sensitive they are to varying kinds of information.

Furthermore, the data regarding reference to non-human animals shows that the association between logophoricity and hän is able to ‘override’ the association between hän and human referents: hän – typically regarded as the human third person pronoun – can be used for animals, even in the absence of personification, if the animal is conceptualized as the character whose thoughts we are being presented with (SELF). This suggests that in addition to taking into account multiple types of information as mentioned above, the system also needs to be able to reflect the relative ranking/weighting of different kinds of information (e.g., being SELF matters more than being (non)human).

Taken as a whole, the data presented here suggest that the referential properties of different anaphoric forms cannot be reduced to a single factor. Our model of anaphora resolution must be flexible enough to incorporate different kinds of information, including prominence, (non)logophoricity, and register variation.

Sources


References