ON THE DEVELOPMENT OF THE DEFINITE ARTICLE *SE* IN SPOKEN FINNISH[1]

Ritva Laury

Department of Linguistics

University of California, Santa Barbara

1. INTRODUCTION

In expressing the central discourse-pragmatic distinction between definite and indefinite reference, languages employ a variety of different means. A speaker of English can use the definite article to indicate that he assumes that the hearer can identify the referent of the noun phrase in question, while speakers of languages that lack articles have to resort to other means to convey this distinction, if they express it at all. Finnish is usually believed to belong to the latter group. Thus Lauri Hakulinen in his 1979 Suomen kielen rakenne ja kehitys states that since Finnish "lacks an expedient formal means of indicating definiteness, such as an article[2]" (510), Finnish instead indicates the distinction between definiteness and indefiniteness through other means, such as case marking, verb agreement and word order. However, as Hakulinen indicates in a footnote, this statement only holds for written Finnish; in dialectal language, he says, there are "clear signs of apparently language-internal article formation" (510).

Hakulinen's remark refers to certain prenominal uses of the demonstrative pronoun *se* in spoken texts collected in the last part of the 19th century. Comparing the data he refers to with records of present-day spoken

Finnish, one can readily see that the development which could be detected a hundred years ago has progressed further along the road to the grammaticization of *se* as an article.

In this paper, I will compare the 19th century spoken data which Hakulinen refers to, to spoken discourse from present-day Finnish. I will show how the demonstrative pronoun *se* is being grammaticized in Finnish into a definite article so that speakers now regularly mark identifiable noun phrases with *se*. I will also show that the diachronic development of the article can best be explained from discourse.

2. THE DATA

My data for this discussion are two narratives from the 1890s, which I will compare with spoken narratives recorded in 1984[3]. The earlier narratives are traditional folktales from two different dialectal areas of Finnish; the later narratives are based on a short film which was shown to undergraduates at the University of Helsinki. Forty-eight narratives were elicited, and six of these narratives were used for this study. These speakers came from various dialectal areas of Finland.

3. *SE* AND DEFINITENESS IN FINNISH GRAMMAR

Finnish grammarians agree that Finnish has three demonstrative pronouns, *tämä*, *tuo*, and *se* (Setälä 1891, Penttilä 1963 and Karlsson 1987). Just like the English *this* and *that*, the Finnish demonstratives can be used both independently and prenominally.

However, it is not entirely clear how these demonstratives should be placed on the proximal/distal scale. While it seems clear that *tämä* is a proximal demonstrative and could be translated into English as 'this', the relationship of the distal demonstratives *tuo* and *se*, both of which could be translated with the English 'that', is not entirely clear. Grammarians disagree about their relative position on the scale, some saying that *se* picks out referents at a medium distance and *tuo* farthest away, while others claim that *tuo* is the medial demonstrative and *se* the distal one; a third position claims that *se* is neutral with respect to distance. (Larjavaara 1985:28). Thus it seems clear that the status of *se* as a member of the Finnish demonstrative pronoun system is different from that of *tämä* and *tuo*.

Some grammarians, such as Setälä, have acknowledged the different status of *se* by saying that *se* is more weakly demonstrative than *tämä* and tuo, and refers to something which "could be pointed out but doesn't need to because it has already been mentioned previously" (1891:76). Penttilä (1963:510) essentially agrees, but adds to Setälä's definition by saying that *se* refers to an object or being which ... has already been brought to attention in some way or another" (1963:510). Most recently, Larjavaara (1985;1990) has suggested that *se* differs from *tämä* and *tuo* in that *se* is addressee-centered while *tämä* and *tuo* are speaker-centered.

Setälä's and Penttilä's comments are most interesting to my discussion. I will show that *se* is developing into a definite article in Finnish. As much recent literature has shown (for example, Chafe 1976 and 1987, Du Bois 1980, and Du Bois and Thompson In progress), formally indicated definiteness in language is closely connected with the speaker's model of the

addressee's consciousness; this brings to mind Larjavaara's category 'addressee-centered'.

According to Chafe (1987), a particular concept, (the idea of an object, event, or property), at a particular point in time, "may be in any one of three different activation states (25)." Information which the speaker assumes to be in the active or focal consciousness of the addressee at the time of speaking, **given** information, is expressed differently from **new** information, which the speaker assumes the addressee not to be conscious of, and **accessible** information, which the addressee is peripherally conscious of. Setälä's comment that *se* refers to something which has already been mentioned seems to be akin to the status of information as given or accessible from previous mention. Penttilä seems to indicate that givenness or accessibility can result from factors other than previous mention.

In my discussion, I will show that the concept of activation states is extremely relevant to the development of the Finnish *se*; a further concept, which Du Bois (1980) calls referentiality and Du Bois and Thompson (In progress) call discourse referential function, the function of a given noun phrase "at the particular sequential point in the discourse at which it is introduced" (23), will also be shown to relevant to this discussion. According to Du Bois and Thompson, **tracking** noun phrases are those noun phrases which are "used to speak about an entity as such, which is conceived of as having continuity of identity (24)." Tracking noun phrases are used by the speaker to track a discourse participant, whereas non-tracking noun phrases are not.

4. ARTICLE VS. DEMONSTRATIVE - IS THERE A DIFFERENCE?

A crucial question for this discussion is the distinction between articles and demonstratives. How exactly are articles different from demonstratives? Can they be clearly distinguished, or are they perhaps endpoints on a scale stretching from clear articles to clear demonstratives, with a 'fuzzy' middle area where the function of a given form cannot be clearly established as either a pure article or a pure demonstrative?

The use of demonstratives in language is closely linked to the concept of deixis, and as many scholars have noted, there are close connections between definiteness and deixis. Levinson suggests that "definiteness may perhaps be an essentially deictic notion" (1983:83), and Lyons claims that anaphoric use of the definite article, personal pronouns and demonstratives is derived from deixis, and that "the presuppositions of existence and uniqueness which logicians commonly associate with the use of the definite article [...] do not distinguish (it) in English from the demonstratives *this* and *that*" (1975:61). Schachter classifies 'demonstrative modifiers' such as *this* and *that* (in their prenominal use) with articles on semantic and syntactic grounds while he acknowledges their semantic and morphological relatedness to the corresponding demonstrative pronouns (1985:40). Similarly, Clark and Marshall point out that "All the world's languages appear to have demonstratives [...], but many do not have definite articles. In these languages, when a definite reference has to be absolutely clear, a demonstrative is used, as in *that woman*" (1981:46).

While synchronic relatedness of deixis and definiteness thus seems fairly

clear, it is of course also well known that in many languages, demonstratives and articles are in fact diachronically related. Demonstrative pronouns are the source of definite articles in the Germanic and Romance languages.

In discussing the development of the Romance definite article, Harris (1977:249) refers to the "over-rigid separation of the traditionally distinct categories of 'demonstrative' and 'definite article'". Harris goes on to say that while it is well known that the Romance definite article is derived from Latin demonstratives, "the exact nature of this change is rarely, if ever, specified in detail; rather, a very general word such as 'weakening' is used to describe the process involved" (ibid).

How and why such weakening should take place is an interesting question. Meillet (1926), with whom the term 'weakening' originated, links the process of weakening and eventual grammaticization of the weakened form to frequency of use. In his operational definition of the difference between demonstratives and articles, Harris (1977:249-250) is clearly based on Meillet; he states that "what is happening is simply that one of the two semantic features taken as defining characteristics of any demonstrative, the marking of proximity, can no longer be conveyed by the original forms, which are left with only the second of their two functions, that of specification." To Harris, then, the crucial distinction between demonstratives and articles is that the former mark proximity while the latter only specify.

Greenberg (1985) discusses the role of iconicity in the development of articles from demonstratives and points out that in such a development, the demonstrative involved is almost always a distance demonstrative. He

suggests that the mapping of perceptual space into discourse time is the motivation for such a development; anaphoric use of the deictic demonstrative, from which article use is presumed to have developed, would constitute a pointing back to an earlier, presumably more distant, mention. Greenberg further notes that in German and Greek, the source of the definite article is a deictic unmarked for distance (1985:275), which "is described as not distinguishing between near and far but involving only an undifferentiated reference to anything not in the immediate vicinity of the speaker" (1985:275). Greenberg goes on to say that "Historically, loss of accent and sometimes phonetic reduction in the change from deixis to anaphora, mirrors the loss of prominence which comes with the change from making known to the mere expression of something as already known, a change from new to old information" (1985:276).

Harris' and Greenberg's comments are extremely interesting in view of the uncertainty among Finnish grammarians about the status of *se* with respect of its place on the proximal-distal scale. *Se* seems to be, just like the source morphemes of the German and Greek articles, unmarked for distance. If the defining characteristic of a demonstrative is the marking of distance, as Harris suggests, *se* is not as much of a demonstrative as *tämä* and *tuo*. Greenberg appears to suggest that in the development of an article from a demonstrative, activation status of the referent plays an important role.

I will show that the Finnish demonstrative *se* has increased in textual frequency as it has approached grammaticization as an article; at the same time, it has weakened phonetically and in its article-like use is always unstressed in pre-nominal position in present-day spoken Finnish. It is not

clear whether the distinction between a demonstrative and an article can be made; instead, there seems to be a gradual development. But two other questions remain to be answered beyond the mere frequency count and phonetic weakening. What provides the causation for increased use of the demonstratives, and when use becomes more frequent, in what kinds of contexts are the weakened uses first encountered?

5. LATE 19TH CENTURY TEXTS

In this section of the paper I will discuss the use of *se* in spoken Finnish approximately a hundred years ago in two texts collected by Salu Latvala in 1890 and 1894 in Satakunta and Savo. The use of *se* is very similar in both of these texts, and in his commentary, Latvala remarks that *se* is used with the "meaning of a definite article" (1899:42) and that *se* is "used as an article" (1895:46). Latvala does not explicate his statements further, other than providing example sentences. What does he mean? What kinds of NPs get articles, and what kinds do not? How frequent is the use of *se* prenominally?

In the Satakunta narrative, 22 of the 145 full NPs in the text, or 15.9%, were preceded by se; in the Savo narrative, 43 of the 366, giving a percentage of 11.2%.

What determines which NPs are preceded by *se* and which are not? The two features which best predict the use of *se* prenominally are accessibility and noteworthiness. Speakers are most likely to mark with *se* those NPs whose referents are accessible from previous discourse context; occasionally, given NPs are also marked with *se*, but not as often as accessible NPs. Out

of the 116 accessible NPs in the two stories, 35%, or 41, were marked with
se, while only 11% of the given NPs were (28 out of 249). New NPs were
never prefaced with *se*. Interestingly, while accessible NPs are simply
prefaced with *se*, the typical pattern for prefacing given NPs with *se* is in a
right dislocation.

Additionally, the NPs which are prefaced with *se* have referents which
are important in some way in the narrative. Thus an NP which is preceded
with *se* is typically an NP whose referent is a central character in the story or
an otherwise crucial prop.

I will illustrate these points with examples from the two narratives.

The Satakunta narrative is a variation on the well-known Bluebeard
theme. The chief of a band of thieves visits a house which has three
daughters. He marries the eldest daughter, brings her into his house, shows
her around but won't show her one room which he instructs her not to visit.
The husband leaves for the woods, and the wife decides to investigate the
house, since this is her house. She goes ahead and looks around.

(1) Isäntä oli antanu hänelle kultasen   omenan   merkiks
    husband had given 3SG[4]-ALL golden-ACC apple-ACC sign-TRANSL
    'The husband had given her a golden apple as a sign'

(2) kun hän mettään   lähti.
    as 3SG forest-ILLAT went
    when he went to the forest.

(3) sitte hän katto sinne   kamariin   se emäntä

    then 3SG looked SE-ILLAT room-ILLAT SE wife

    Then she looked into the room, the wife.

The two uses of *se* in the excerpt above are the first ones in the narrative. This example illustrates the use of *se* with both accessible and given referents. *Kamari* 'room', is accessible from prior mention while *emäntä* 'wife' is given. The wife is referred to twice within the same sentence, first with the pronoun *hän* and then with the right dislocated *se emäntä*, while *kamari* is simply prefaced with *se*. Both NPs are highly noteworthy at this point in the narrative. Compare the marking of these NPs to the other NPs in this example; *isäntä* 'husband' is a given referent referred to with a full NP, and not a noteworthy participant in this part of the story, so it doesn't get marked with *se*. *Omena* 'apple', is a new referent so it can not be marked with *se*; and *mettä* 'forest' is accessible, but it is not prominent, and so does not get marked with *se*.

Note also the clustering of se-marked NPs at a crucial point in the narrative. The wife's looking into the room is pivotal at this point in the story, as it is the cause of her imminent demise at the hands of her husband. Clauses 4 and 5 are backgrounded with respect to clause 6 which is foregrounded. Thus se-marked NPs don't only occur with noteworthy participants in a narrative, but they also seem to cluster in foregrounded clauses.

The narrative continues. The room (pronominalized) is full of plates with human heads with blood dripping into a big vat. The wife starts to worry about what will happen to her in such a place.

(4) Sitte hänen se kultanen omena putos sinne[5] ammeeseen.

then 3SG-GEN SE golden apple fell SE.LOC-ILLAT vat-ILLAT

then her golden apple fell into the vat.

Sentence 4 is another example of accessible NPs being prefaced by *se* when they become highly topical at a crucial point in the narrative. Especially striking is the use of *se* with a possessive pronoun.

When the apple falls into the vat, the gold comes off it. When the husband sees this, he kills his wife. A second daughter he marries from the same house meets with the same fate, but the third daughter is smarter. She manages to avoid detection after she sees the secret room, and instead takes her husband on a trip to her parents' house. As evidence, she takes along her sisters' cut off heads from the secret room in a chest. Another crucial point in the narrative takes place when, after an exciting journey, the daughter and her husband arrive at her home. The wife instructs her parents to get her husband drunk. The parents comply with their daughter's request.

(5) Hän näytti sitte vanhemmilles   siältä   kistusta

3SG   showed then parents-ALLAT-her SE.LOC-ABL chest-ELAT

She showed then to her parents from the chest

ne   sisartes   päät.

SE.PL sisters-GEN-her heads

the heads of her sisters'.

Again, the accessible NPs *kistu* 'chest' and *sisartes päät* 'her sisters'

heads' are prefaced with se. They are extremely noteworthy in the narrative because they allow the wife to prove her husband's villainy. At the same time, the clause is highly foregrounded as it marks a turning point in the narrative in that the husband's villainy is now revealed and the family can get its revenge.

The other narrative from the Savo area is less gruesome but similar use of *se* can be illustrated from this narrative also. In this story, apples are disappearing from a king's orchard. The king has three sons, who take turns watching the orchard at night. The two elder sons see nothing; then,

(6) sen   nuorimman   poejan - sen   nim ol   Junnu -
SE-GEN youngest-GEN boy-GEN   3SG-GEN name was Junnu
the youngest boy - his name was Junnu -

rupes   tekemään miel' vahtiin
started making   mind watch-ILLAT
started to want to go on watch.

Although the youngest son is accessible from the fact that the king had three sons, he has not been mentioned previously. When the elder sons were introduced into the narrative one by one, the NPs which introduced them were not prefaced with se; the first introduction of the youngest son in (6) is prefaced with *se* because his role in the narrative is more important than that of the other sons. He is the protagonist.

Junnu manages to stay awake. At midnight, a bird flies up and tries to

steal an apple. Junnu tries to catch it but only manages to hold on to a tailfeather. In the morning, when he returns home, he's asked whether he saw anything. He says yes, but the others call him a liar because they hadn't seen anything. He tells them to pull the curtains. The curtains are pulled.

(7) se   sitte otti sen   hööhenen   lakkaristaan
     3SG then took SE-ACC feather-ACC pocket-ELAT-his
     He then took the feather from his pocket

     niin se valas   sen   huoneen
     so   3SG lit.up SE-ACC room-ACC
     and it lit up se room.

Both NPs marked with *se* are accessible; *hööhen* 'feather' is accessible from prior mention. It is also a crucial prop because it is with the feather that Junnu proves the existence with a golden bird which then becomes the object of a search which constitutes the main portion of the narrative. *Huone* 'room' is also accessible, but from context; the reference to the curtains makes the presence of the participants in a room clear. However, the room is not in any other way crucial, noteworthy or topical. I hypothesize that it is the fact that the room is an accessible NP in a foregrounded clause which accounts for its being marked with se. It may be that it is these usages, where *se* is associated with only accessibility and not noteworthiness, that are the pathway to full grammaticization of *se* as an article.

Is it clear whether *se* is being used in these narratives as an article or a demonstrative?   As Kirsner (1979) and Mithun (1987) have shown,

demonstratives can be used both in languages which have articles and languages which do not have them in important discourse-organizational functions without having to be classified as articles.

Kirsner's study of the proximal and distal demonstrative use in modern Dutch does not directly address the demonstrative/article distinction, but he does show that the concept of distance is not crucial to the use of the Dutch equivalents of this and that, but that their use is instead connected with 1) noteworthiness, the tendency of a speaker to "direct attention strongest to entities that he, the speaker, is most interested in talking about" 2) givenness; the tendency, after Chafe (1974) "to direct the hearer's attention strongest to entities that are not given" and 3) foregrounding; the speaker's tendency for "drawing the appropriate amount of attention to the noun's referent, so that strong urging of the hearer to find it will be coupled with devices for foregrounding the noun in question."(1979:360)

Since *se* is used in these texts primarily for foregrounding noteworthy referents which are accessible, could it be that it is still a demonstrative, and not an article? One reason to call *se* at least a developing article is that its use contrasts very distinctly with the other two forms which Finnish grammars call demonstratives, *tämä* and *tuo*. In the two texts I have discussed above, *tämä* and *tuo* appear only in direct quotes. Mayes (1991) shows that deictic terms in quoted speech are chosen "as if the quote were the original utterance and the speaker were the original speaker (346)." Mayes suggests that the "highlighting" or "theatrical" quality attributed to quotes "is, for the most part, due to the fact that the deictic center of a direct quote is that of the original event (346)." The demonstratives *tämä* and *tuo* in the narratives I

examined serve this function of deictically marking the quotes as pronounced not at the time of telling, but at the time the (imagined) original event. This is quite different from the function of *se*, which has to do with the status of the NP within the text itself, in the universe of discourse instead of the story world.

Below are examples of the use of *tämä* and *tuo* in quotes in the Savo and Satakunta stories:

(8) sano, jott' "älä   mene tuanne   kamariin
    said COMPL don't go that-ILLAT room-ILLAT
    (he) said, don't go into that room

(9) "juattakaa   nyt tämä minun miäheni
    drink-CAUS-IMPERAT now this my   husband-my
    make this husband of mine drink

(10) tuossahan   se on tuo toinen pelar
     that-INESS-PTCL 3SG is that other player
     there he is, that other player

One might even speculate that the usage in quotes is the original pathway for the association of demonstratives with prominent referents. Mayes shows that quotations function to highlight main points in a narrative (1991:358). I have suggested that *se*-marked accessible referents typically appear in highlighted portions of narratives also. Certainly the function of *se*-marking of referents in the two texts I have examined from the 19th century

is to iconically highlight referents which the narrator wants the hearer to recall and pay attention to.

It is interesting to note that Christophersen in his discussion of the development of the English definite article says that "the extended use of the resumptive pronoun is only found with the chief characters and localities and things belonging to them" (1939:87). By 'resumptive' Christophersen appears to mean use for accessible referents. Thus the English article in its early stages appears to have been used in much the same way as the Finnish se, to highlight accessible, prominent referents.

6. THE USE OF SE IN PRESENT-DAY SPOKEN FINNISH

An examination of spoken narratives by speakers of modern spoken Finnish shows that *se* is well on its way to being grammaticized as a definite article. Speakers consistently preface NPs with *se* when they have reason to expect that the hearer can identify the referent of that NP.

In comparison with the earlier narratives, there has been a dramatic increase in the use of *se*, and speakers now regularly preface given as well as accessible referents with *se*. From a total of 463 full NPs tabulated for the modern data, 233, or 50.3%, were preceded by *se*; this percentage was only 11.2% for the Savo narrative and 15.9% for the Satakunta narrative in the earlier data.

In the Pear story data, given NPs are marked with *se* even more frequently than accessible NPs; 52% of the accessible NPs are marked with *se*, and 69%

of the given NPs are, compared to 35% for accessible NPs and only 11% for given NPs in the early narratives. Given and accessible concepts share the feature of identifiability, and in modern spoken Finnish, *se* is a marker of identifiability. It is also interesting to note that out of all full NPs referring to humans in the modern data, 85.3% were preceded by either *se* or *tämä*.

Tämä, which in the earlier narratives was restricted to quotations, has also increased in frequency and taken over the highlighting function I attributed to *se* in the earlier data. *Tämä* was only used once prenominally in the Satakunta narrative out of 145 full NPs, giving a figure of 0.7%; it was slightly more frequent in the Savo narrative, where it was used six times. Thus 1.6% of the full NPs in the Savo narrative were preceded by *tämä*. This same figure for the Pear narratives is 35%.

The narratives in this section are based on a short film. The film opens with a shot of a tree from which an older man is picking pears. Another man leading a goat walks by the tree. Then a young boy rides on a bicycle. Unbeknownst to the man, he steals a basket of pears. He then passes a girl on his way and falls down, and the pears spill from the basket on his bicycle. A group of three boys come to help him gather the pears; as a reward, he gives each boy a pear. The three boys walk past the tree eating the pears just as the pear-picker has discovered the theft of his basket of pears.

Below are examples which illustrate the use of *se* in the Pear narratives[6].

(11) ... 'keräs,

    picked

    was picking

(12) ... ^päärynöitä.

    pears-PARTIT

    pears.

(13) .. ^Puusta.

    tree-ELAT

    From a tree.

(14) ... 'Kiipes ^tikapuita   pitki sinne,

    climbed ladders-PART along SE.LOC-ILLAT

    (He) climbed along the ladder into the

(15) ... (1.2) ^puuhun,

    tree-ILLAT

    tree

(16) ... ^repimään   niitä   päärynöitä.

    tear-3INF-ILLAT SE.PL-PART   pears-PART

    to tear out the pears.

(17) .. Mä en   tiedä päärynöistä paljoo mut,

    I don't know pears-ELAT much   but

    I don't know much about pears but,

(18) must   ^tuntuu et,

1SG-ELAT feels COMPL

I think that,

(19) ne   'päärynät oli 'kaikki ^raa=koja.

SE.PL pears   were all   raw-PL-PART

the pears were all green.

In lines 12 and 13, *päärynöitä* 'pears' and *puusta* 'tree'(in the elative case) are new, and therefore not preceded by *se*; these are the first two uses of these NPs in this particular Pear story. The *puuhun* 'tree'(illative) on line 15 and *päärynöitä* 'pears' on line 16 are given, and preceded by the plural form of *se*.

The next use of *päärynöistä* 'pears' (elative) on line 17 is another use of the same NP, but not preceded by *se*. This is because this use of the NP is what Du Bois and Thompson (1991:27) call non-tracking. The speaker here is not referring to any particular pears whose identity would be tracked in the narrative in the same way as the pears which the man picks are. Du Bois and Thompson mention that one difference between tracking and non-tracking NPs is that the latter often do not take determiners. This is because "where a NP is not being used for the prototypical function for which NPs exist, namely to track participants, it sheds many of the grammatical trappings which are crucial for participant tracking, but which are not necessary when the NP is not serving to track participants" (1991:38). Note that when the speaker returns to tracking the pears in the story on line 19, he again prefaces

the NP with *se*.

Another pattern which is quite noticeable in the Finnish Pear stories, illustrated in lines 14 and 15 above, is the division of *se*-prefaced NPs by an intonation unit boundary. *Se* often occurs at the end of one intonation unit, and the noun phrase follows, in a separate intonation unit. One could hypothesize that this pattern originates when the speaker intends to pronominalize the referent, and then decides the referent is not quite given enough for the addressee to be able to decode it from only a pronoun. In any case, the pause before line 15 can not be interpreted as a hesitation as to the NP with which to code the referent as the speaker has just used the noun puu on line 13. Below is another example of this usage from a different Pear story.

(20) ... se=,
    SE
    the=,

(21) .. ^mies ei huomannu.
    man NEG notice
    man didn't notice.

The NP on lines 20-21 refers to the pearpicker. The referent is accessible, not given at this point; it has been mentioned previously 12 intonation units back; the previous mention was a pronoun. Very possibly the speaker here has some doubts about whether the addressee could retrieve the referent of the pronoun *se* if it had occurred by itself, particularly as there had been an

intervening human referent, the boy on the bicycle. It is very possible that this pattern may have grammaticized into a way of marking accessible referents. Furthermore, note how dividing an NP onto two intonation units would tend to highlight a referent and make it more prominent. Thus this pattern would contribute to the highlighting function of *se* marked NPs, which I have noted previously.

Another example contrasts the use of *se* with the use of *tämä*. The speaker has just introduced the goat which goes by the pear tree in the beginning of the movie.

(22) ... Yks[7] ^mies toi   sitä,
         one man brought 3SG-PART
         A man brought it

(23) ... ^narun   päästä.
         string-GEN end-ELAT
         at the end of a string.

(24) ... SEN   ^puun   ohitte.
         SE-GEN tree-GEN past
         Past the tree

(25) .. Mistä   tämä ^mies keräs   päärynöitä ja.
        which-ELAT TÄMÄ man gathered pears-PART and.
        From which this man gathered pears.

Mies 'man', mentioned on line 25 is accessible from prior mention in the discourse from 20 intonation units back, as is the tree. Why does the NP referring to the man get prefaced with *tämä* while the NP referring to the pear tree is prefaced with *se*? As I have mentioned previously, the data suggests that *tämä* in the modern data has taken over the highlighting function that *se* had in the older data. Possibly the speaker is highlighting the noun *mies* 'man' because there has been an intervening potentially coreferential mention *yks mies* 'a man' on line 22. But in the immediately preceding discourse, the speaker is discussing the scenery in the film and trees are mentioned twice in this part of the story; yet *puu* 'tree' on line 24 is not highlighted by *tämä*. Why?

In his discussion of the discourse functions of the Dutch demonstratives, Kirsner shows that while Dutch has a definite article, demonstrative pronouns are used when a speaker wants to highlight a referent. He further shows that human referents are more likely to be marked with a demonstrative than are other NPs (Kirsner 1979:360). As mentioned above, it appears from the Finnish data also that human NPs are more likely than other NPs to be marked with *tämä*.

The changed function of *tämä* also underscores the centrality of the highlighting function demonstratives typically perform in language. Recall that in the narrative data from the 19th century, the demonstratives *tämä* and *tuo* had a restricted distribution; they were only used in quotations. This indicates that at that time, these demonstratives still were used in the more concrete, or basic, deictic use typical of demonstratives. The highlighting function performed by *se* earlier is apparently so important that when *se*

became grammaticized as a marker of identifiability, *tämä* took over its earlier function.

8. CONCLUSION

In this paper I have provided an explanation for the development of the demonstrative pronoun *se* in spoken Finnish. I have shown that the demonstrative pronoun *se* has advance further to becoming grammaticized in spoken Finnish as a definite article. From presumably having earlier functioned as a pure demonstrative which picks out new referents by pointing out their concrete spatial location, by the 19th century *se* had come to have a highlighting function in discourse; it was used with prominent, accessible referents. Accessible and given concepts share the feature that they are identifiable, and languages which have articles, such as English, mark both accessible and given NPs as definite. Through becoming associated with the feature of identifiability, *se* has undergone reanalysis and is being grammaticized as a definite article in spoken Finnish. After *se* became a marker of identifiability, it lost its function of marking NPs as prominent, and this function was taken over by *tämä*.

The grammaticization process discussed here involves the kind of semantic change described in Traugott (forthcoming) as "the shift from meanings grounded in more or less objectively identifiable extralinguistic situations to meanings grounded in text making". Although grammaticization has traditionally been defined as "a process where lexical items in the course of time acquire a new status as grammatical, morpho-syntactic forms"(ibid), it is thus possible to include within the concept of grammaticization changes

in the functions of grammatical morphemes. The motivation for this change was shown to be the need for speakers to signal to their addressees the information flow status of the referents of the linguistics signals they use.

REFERENCES

Chafe, Wallace. 1976. Givenness, contrastiveness, definiteness, subjects and topics. Subject and Topic, ed. by Charles Li. New York: Academic Press.

Chafe, Wallace. 1987. Cognitive constraints on information flow. Coherence and Grounding in Discourse, ed. by Russell S. Tomlin. Amsterdam: John Benjamins.

Christophersen, Paul. 1939. The Articles. Copenhagen: Einar Munksgaard.

Clark, H.H. and C.R. Marshall. 1981. Definite reference and mutual knowledge. Elements of Discourse Understanding, ed. by A.K. Joshi, B.L. Webber, I.A. Sag. Cambridge: Cambridge University Press.

Du Bois, John W. 1980. The trace of identity in discourse. The Pear Stories, ed. by Wallace Chafe. Norwood: Ablex.

Du Bois, John W., Stephan Schuetze-Coburn, Danae Paolino and Susanna Cumming. 1990. Discourse Transcription. Santa Barbara: UCSB Dept. of Linguistics.

Du Bois, John and Sandra Thompson. In progress. Dimensions of a theory of information flow. Ms., UCSB, Dept. of Linguistics.

Greenberg, Joseph H. 1985. Some iconic relatonships among place, time, and discourse deixis. Iconicity in Syntax, ed. by John Haiman. Amsterdam: John

Benjamins.

Hakulinen, Lauri. 1979. Suomen kielen rakenne ja kehitys. Helsinki: Otava.

Harris, Martin. 1977. 'Demonstratives', 'articles', and 'third person pronouns' in French: changes in progress. Zeitschrift für Romanische Philologie 93:249-261.

Karlsson, Fred. 1987. Suomen peruskielioppi. Pieksämäki: Suomalaisen Kirjallisuuden Seura.

Kirsner, Robert. 1979. Deixis in discourse: An exploratory quantitative study of the modern Dutch demonstrative adjectives. Discourse and Syntax. Ed. by Talmy Givon. (Syntax and semantics, 12) New York: Academic Press.

Larjavaara, Matti. 1985. Suomen demonstratiivisysteemin rakenne. Sananjalka 27:15-31.

Larjavaara, Matti. 1990. Suomen deiksis. Helsinki: Suomalaisen Kirjallisuuden Seura.

Latvala, Salu. 1895. Lauseopillisia havaintoja Luoteis-Satakunnan kansankielestä. Suomi 12:1-89

Latvala, Salu. 1899. Lauseopillisia muistiinpanoja Pohjois-Savon murteesta. Suomi 17:1-90.

Levinson, Stephen C. 1983. Pragmatics. Cambridge: Cambridge University Press.

Lyons, John. 1975. Deixis as the source of reference. Formal Semantics of Natural Language, ed. by Edward L. Keenan. Cambridge: Cambridge University Press.

Lyons, John. 1977. Semantics. Volume 2. Cambridge: Cambridge University Press.

Mayes, Patricia. 1991. Quotation in spoken English. Studies in Language 14-2.325-363.

Meillet, A. 1926. Linguistique Historique et Linguistique Générale. Paris: Librairie Ancienne.

Mithun, Marianne. 1987. The Grammatical Nature and Discourse Power of Demonstratives. BLS 13:184-194.

Penttilä, Aarni. 1963. Suomen kielioppi. Porvoo: WSOY.

Schachter, Paul. 1985. Parts-of-speech systems. Language Typology and Syntactic Description, ed. by Timothy Shopen. Vol. 1. Cambridge: Cambridge University Press.

Setälä, E.N. 1891. Suomen kielen lauseoppi. 3. muutettu painos. Helsinki.

1.I have greatly benefited from discussions with Sandy Thompson, Wally Chafe, Jack DuBois and Bill Ashby while this work was in progress. I also thank Charles Li, Auli Hakulinen, Marja-Liisa Helasvuo and the SKY anonymous reviewer for their insightful and helpful comments. I am, of course, alone responsible for the shortcomings of this work. I also thank the American Scandinavian Foundation for their financial support while I was preparing this article for publication.

2. Here, as in all direct citations from my Finnish sources, the translation is mine. Needless to say, I have tried to reflect as accurately as possible the Finnish-language original in wording and intent.

3. It should be noted that because the older narratives were, naturally, not tape-recorded but rather done by hand, the transcription may not be entirely accurate. Traditional narratives may also be crystallized from many retellings, while the modern narratives discussed here were spontaneously produduced, based on a film the narrators had just viewed, and recorded at the first telling. Since the corpora are thus not fully comparable, the results should also be considered tentative.

4. Although the Satakunta dialect makes a distinction between singular pronouns referring to human and nonhuman referents, I have coded all occurrences of independent se and hän as 3SG in the morphological glosses. Naturally, the appropriate English pronouns are used in the free glosses. Prenominal uses of se are glossed as SE in the morphological glosses to distinguish them from the independent uses of se.

5. I consider the locative-adverbial forms of se functionally comparable to other prenominal use of se, and my glossing here reflects that fact.

6.The transcription system used for the examples in this section is tthe one described in Du Bois et al. 1990.

7.The numeral yksi 'one' is showing signs of becoming an indefinite aarticle in Finnish in a development analogous to English. I will nnot discuss this development here.