Preliminary Analysis of Free Direct Speech in *Se tapahtui täällä* by Raija Siekkinen

Project Topic and Humanities Research Questions

This project is a preliminary analysis of free direct speech (henceforth FDS) in a literary text which is also a work of art. FDS is a mode for presenting characters' speech in fictional prose – in a direct and free manner. FDS is a direct, untagged quote, free from an introductory verb or clause and typographical signs (brackets, colon and dash). FDS has been widely used as a technique by different authors to mimic dialogue. Segments that include FDS tend to increase the illusion of "spoken" fictional dialogues. The co(n)text will help the reader to interpret what is said and usually by whom. The terminology of reported speech and thought used in this project is established by Leech & Short (1981), McHale (1978) and Cohn (1978).

TERM	ABBREVIATION
Direct speech	DS
Indirect speech	ID
Free indirect speech	FIS
Free direct speech	FDS
Direct thought	DT
Indirect thought	IT
Free indirect thought	FIT
Free direct thought	FDT
Reported speech/Diegetic summary	RS
Narrated perception	NP
Psycho-narration	PN

The humanities research questions concerning this project are:

- How widely and frequently is FDS used?
- What other forms of reported speech and thought precede FDS or follow it?
- In what kind of textual context is FDS likely to appear?

With the help of computational readymade tools, the aim of this project is to find out the quantity and distribution of FDS in the corpus. The tools also indicate the amount of the other forms of reported speech and thought.

Data & Data Processing

The data consists of a textual corpus extracted from the novel *Se tapahtui täällä* (1999) by Raija Siekkinen. The novel is written in Finnish and contains 38 chapters that are divided into three sections. For this analysis, I chose the chapters 12, 18 and 20 because they manifest a large number of occurrences in direct and free direct speech. First, I extracted the data by copying it

manually to a Word-file (the original text contains some typos), secondly, I analysed the forms of different reported speech and thought and tagged the segments with the abbreviations explained above. The possible biases in my analysis could be due to the fact that defining different forms of reported speech and thought is an interpretative choice.

After annotating and tagging the different forms of reported speech and thought, I pasted the texts to the web page Voyant Tools. It is a web-based text reading and analysis environment. After pasting the data, Voyant Tools provides you with different tools to count words and their frequencies and to visualize the distribution, for example. Before analysing the most frequent items, I made list a of stopwords in Finnish (*ja, ei, ole, oli, on, hän, niin*). In order to extract different type of information and visualization using the tools, I edited the options to show only the different forms of RS (by adding items on the search box beneath each tool). Initially, I analysed each chapter separately until I discovered how to paste them all into the Voyant interface.

Data Analysis & Results

Altogether, the Figure 1 (Terms) shows the number of each form of reported speech and thought in all three chapters. The most frequent form of RS is DS (68 counts), followed by FDS (35) and PN (13). Since all three chapter include a lot of dialogue, the high number of direct speech was expected.

Terms

	Term	Count
1	ds	68
2	fds	35
3	pn*	13
4	dt*	7
5	rs*	7
6	it*	5
7	fit*	3
8	id*	2
9	np*	2
10	fdt*	0
11	fis*	0

Figure 1. The number of different forms of reported speech and thought.

Figure 2 (Bubblelines tool) displays and visualizes the frequency and distribution of the forms of reported speech and thought. Each chapter of the corpus is represented as horizontal line and divided into segments of equal length. Each term (DS, FDS, PN, RS, DT) is displayed as a bubble with the size of the bubble showing how frequent the term is in the corresponding segment of text. Generally, the larger the bubble the more frequent the term is.

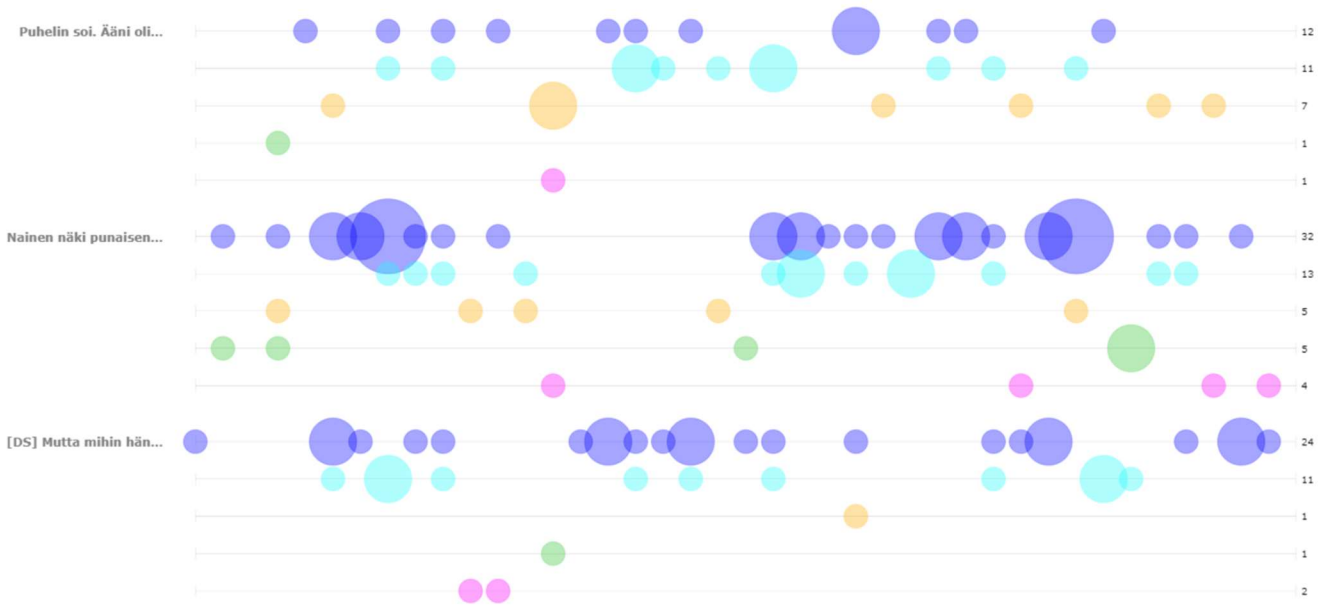

Figure 2. Frequency of DS (blue), FDS (turquoise), PN (yellow), RS (green) and DT (pink).

From the Figure 2, it is possible to see that the most frequent item (DS) and the second frequent item (FDS) correlate and collocate with each other. The larger bubbles indicate that in that particular section of the text the corresponding form of RS is frequently used. The Voyant tools do not take into consideration the length of a segment in a certain mode of reported speech and thought, they merely indicate how often a certain form appears and what other forms precede or follow it.

Figure 3 shows the same five most frequent items on top of each other. From this figure, we can conclude that DS and FDS co-occur most often.

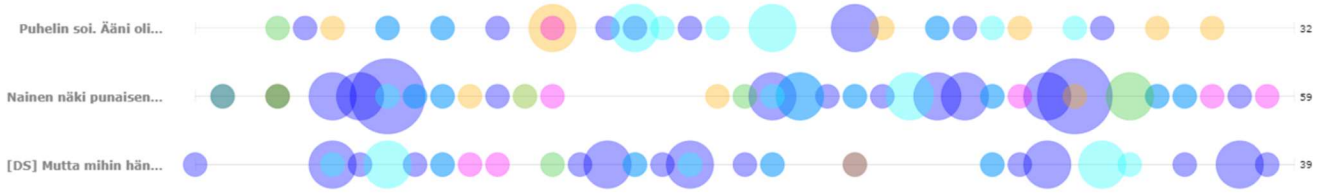

Figure 3. Frequency of DS (blue), FDS (turquoise), PN (yellow), RS (green) and DT (pink).

Figures 4, 5 and 6 display the relative frequencies of the five most frequent forms of reported speech and thought. The figures were made with Voyant Trends tool which shows a line graph depicting the distribution of a term's occurrence across the corpus.

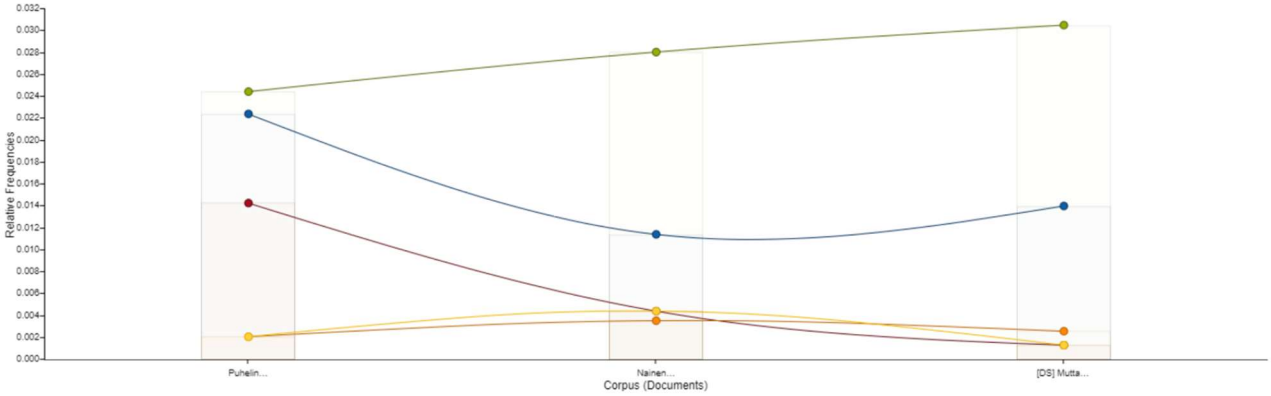

Figure 4. Relative frequencies of DS (green), FDS (blue), PN (red), RS (yellow) and DT (orange).

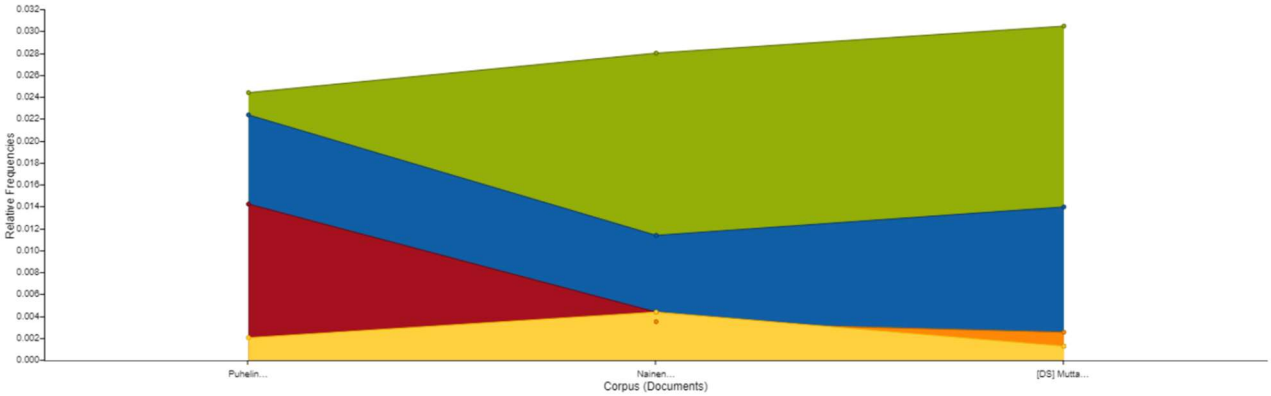

Figure 5. Relative frequencies of DS (green), FDS (blue), PN (red), RS (yellow) and DT (orange).

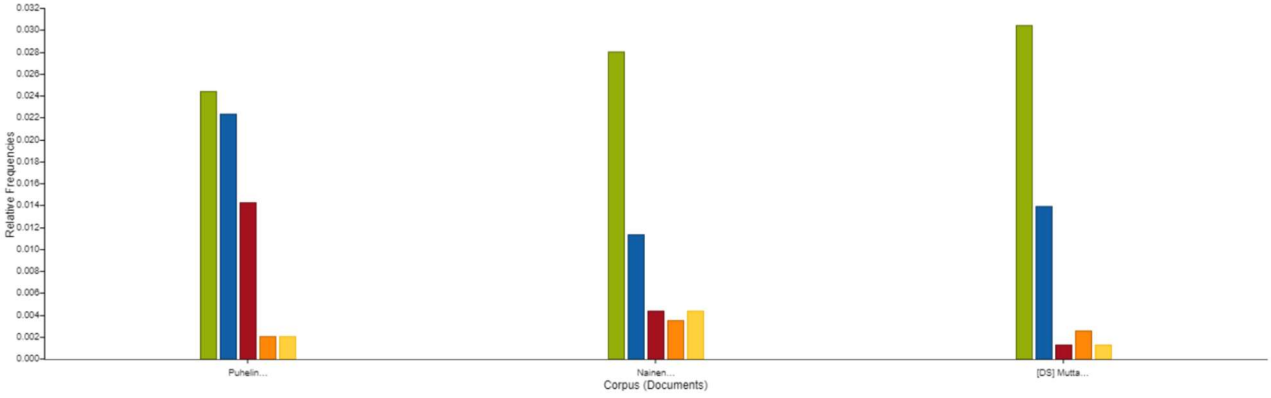

Figure 6. Relative frequencies of DS (green), FDS (blue), PN (red), RS (yellow) and DT (orange).

Figure 4 (Trends line and stacked bar), Figure 5 (Trends area) and Figure 6 (Trends column) all indicate that the most frequent form of reported speech and thought in the corpus are DS and FDS.

This preliminary analysis of FDS has shown that FDS is a quite widely used mode of reported speech and co-occurs often with DS. In order to have precise numeric information and percentages of the FDS, I will have to use other computational methods and programming which enable me to count all the words that were represented in that mode. However, this very simple analysis indicates that FDS is connected especially to DS and mostly occurs in segments that contain dialogue.

References

Cohn, D. (1978). *Transparent minds: narrative modes for presenting consciousness in fiction*. Princeton University Press.

Leech, G., & Short, M. (1981). *Style in fiction: a linguistic introduction to English fictional prose.* Longman.

McHale, B. (1978). "Free indirect discourse: a survey of recent accounts." *Poetics and Theory of Literature* 3. 259–287.

Corpus

Siekkinen, R. (1999). *Se tapahtui täällä*. Otava.