jcls. Journal of Computational Literary Studies

OPEN ACCESS

Article

Genre and the Distribution of Dialogue

Direct Speech and Genre in 500 Contemporary Novels in Swedish

Karl Berglund[1]
Mats Dahllöf[2]

1. Department of Literature and Rhetoric, Uppsala University, Uppsala, Sweden.
2. Department of Linguistics and Philology, Uppsala University, Uppsala, Sweden.

Citation
Karl Berglund and Mats Dahllöf (2026). "Genre and the Distribution of Dialogue. Direct Speech and Genre in 500 Contemporary Novels in Swedish". In: *Journal of Computational Literary Studies* 5 (1). 10.48694/jcls.4284

Date published 2026-09-24
Date accepted 2026-06-05
Date received 2025-05-28

Keywords
dialogue, direct speech, computational narratology, contemporary fiction, Swedish fiction, genre fiction

License
CC BY 4.0

Reviewers
Annelen Brunner, Thierry Poibeau

Note
This paper has been submitted to the journal-only track of JCLS.

Abstract. In this article, we analyse the use of direct speech dialogue (DSD) in contemporary prose fiction. The study is based on a corpus of roughly 500 commercially and/or critically successful works of fiction in Swedish from the period 2004–2020. Three aspects of DSD are in focus: 1) the relative amount of DSD; 2) the distribution of DSD within narratives; and 3) the average length of DSD insets. It is shown that the use of DSD follows clear trends but also that it is used in diverse ways. DSD is highly dependent on genre. In particular, we find differences between literary fiction (less dialogue, but more varied) and popular genre fiction (more dialogue, but less varied). Another result is that the length of DSD insets depends on authors to a large extent. Finally, we propose a scheme for how DSD and genre in contemporary fiction can be characterised in quantitative terms.

1. Introduction

Dialogue is such a ubiquitous device in modern fiction that it might seem like one of the least interesting items in the writer's toolbox. Still, as Stephen King notes, there is nothing more disturbing to a reader than poorly written dialogue – "it jags on the ear like a badly tuned musical instrument" (King 2020, 213). For King, dialogue should aim for realism. Stay true to your characters, he encourages the presumptive writer:

> It's dialogue that gives your cast their voices, and is crucial in defining their characters – only what people do tells us more about what they're really like, and talk is sneaky: What people say often conveys their character to others in ways of which they – the speakers – are completely unaware. (King 2020, 210)

It is safe to say that how dialogue is designed and the purposes it serves in many ways set it apart from the narrative in which it is embedded. Still, dialogue is not only used in the way advocated by King.

Dialogue serves a range of purposes in prose fiction: It "brings characters to life through their speech" (King 2020, 187); it moves the plot forward and allows characters to come forward in their own voices. Dialogue can also be a tool for creating a sense of hard-boiled and journalistic realism, as well as providing information which seems more relevant for the reader than for the characters of the story.

Despite this, the use of dialogue as a storytelling device has not often been studied using quantitative methods. While there is a body of scholarship addressing how specific authors use dialogue, less is known about common patterns and ways of deviating from those patterns.

The purpose of the present article, then, is twofold:

1. to outline a quantitative empirical methodology for studying how dialogue is used in contemporary prose fiction, centred on answering questions such as: On average, how much dialogue do contemporary novels contain? How does it differ between genres? How is dialogue distributed? For how long are characters speaking?
2. to use this approach to challenge common assumptions about the relationship between presentation of dialogue and genres.

One such assumption is that the distinction between popular genre fiction and prestige literary fiction is partly visible in how authors use dialogue. As Ken Gelder notes, literary fiction "doesn't need a story or a plot, but popular fiction couldn't function without one"; most contemporary popular genre fiction is "built around plot, action, 'scenarios', character conflict and dialogue" (Gelder 2004, 19, 28). Another assumption is that certain popular genres use dialogue differently. For instance, Kim Wilkins et al. claim that crime fiction is associated with "fast-paced dialogue," whereas romance is instead said to be characterised as "largely introspective, focused on the protagonist's thoughts and feelings." (Wilkins et al. 2022, 164)

The present work relies on a methodological approach which has been labelled "quantitative formalism" (Allison et al. 2011). Its emphasis is on analysis of formal conventions in literature by means of quantitative measurements of various textual features. The study also makes use of "computational narratology" (Mani 2013; Meister 2003; Piper et al. 2021). To be more specific, we investigate the dialogue in 500 commercially and/or critically successful works of contemporary prose fiction in Swedish, some translated from other languages.

Three aspects of dialogue are studied in depth:

1. the relative amount of dialogue;
2. the distribution of dialogue within narratives; and
3. the length of the dialogue insets.

By analysing these three parameters both on the cohort level and on the level of individual novels, the study sheds light on the roles of dialogue in contemporary fiction, with special attention paid to how this relates to genre.

We set out by sketching the background for analysing dialogue within a narrative frame, discussing what has been done in previous research, and how we theoretically understand narration and dialogue. We then move on to discuss methodological choices and the corpus of novels we use. In the analysis, one section each is dedicated to the three aspects of dialogue in focus: share, distribution, and length. In these, we move from patterns on the general level to zooming in on outlier examples. In the conclusion

we discuss the wider implications of the framework we have proposed for quantitatively analysing how dialogue and genre are related in contemporary fiction.

2. Literary Narration and Dialogue

2.1 Previous Research

As James Phelan (2005, 2011) notes, classical narratology is not very interested in dialogue. It has been focused on the narration of the main narrator, placing dialogue in a secondary position. In Seymour Chatman's famous "narrative communication model" (Chatman 1978, 151), dialogue is absent. According to Gérard Genette, "[u]p to the end of the nineteenth century, the novelistic scene is conceived, fairly piteously, as a pale copy of the dramatic scene: mimesis at two degrees, imitation of imitation" (Genette 1980, 173), suggesting that dialogue scenes were regarded as foreign or even hostile to the true novel form (Lindemann 2019, 109–111).[1]

This lack of interest might be explained by the notion that character-character conversation is a rather simplistic way of telling a story that highlights plot in a seemingly realistic way (Thomas 2012). It is also a mode of storytelling prominent in popular genre fiction relying on narrative techniques more typical of film and journalism than of belletristic literature. It is not a coincidence that Genette's main example in *Narrative Discourse* (Genette 1980) is Marcel Proust's *À la recherche du temps perdu* (1913–1927), a high-modernist classic that experiments with the narrative form in a variety of ways. Bronwen Thomas argues that Mikhail Bakthin's understanding of direct speech dialogue "as little more than a masquerade or puppet show" (Thomas 2012, 4) does not give us any good reasons not to study dialogue in more depth. Here, we follow Thomas in taking these 'puppet shows' to be important components in the structure and composition of literary works.

There is a vast body of qualitative literary research analysing specific works of fiction where direct speech dialogue is prominent or used in some aesthetically conscious way, and which problematises the often-assumed relationship between dialogue and realism exemplified by Stephen King in this article's introduction (Björck 1953; Fludernik 1993; Herman 2006; Kacandes 2001; Lindemann 2019; Phelan 2005, 2011; Sternberg 1982; Thomas 2012). However, such studies seldom discuss the well-established praxes of dialogue use in popular genre fiction. The more empirically-oriented studies investigating the role of dialogue focus on either linguistic-lexical aspects (McIntyre 2010; Vishnubhotla et al. 2019), historical change (Håkansson and Östman 2018; Jannidis et al. 2018; Muzny et al. 2017), or dialogue share as a potential predictor for fiction (Piper 2018).

We believe this study to be the first that based on a larger corpus investigates how character-character direct speech dialogue functions as a device in contemporary prose fiction.

1. Also in more recent handbooks on narratology, such as Mieke Bal's *Narratology* (2017, 4th ed.), analysis of dialogue in terms of direct speech is only addressed occasionally (Bal 2017, 37, 46–50).

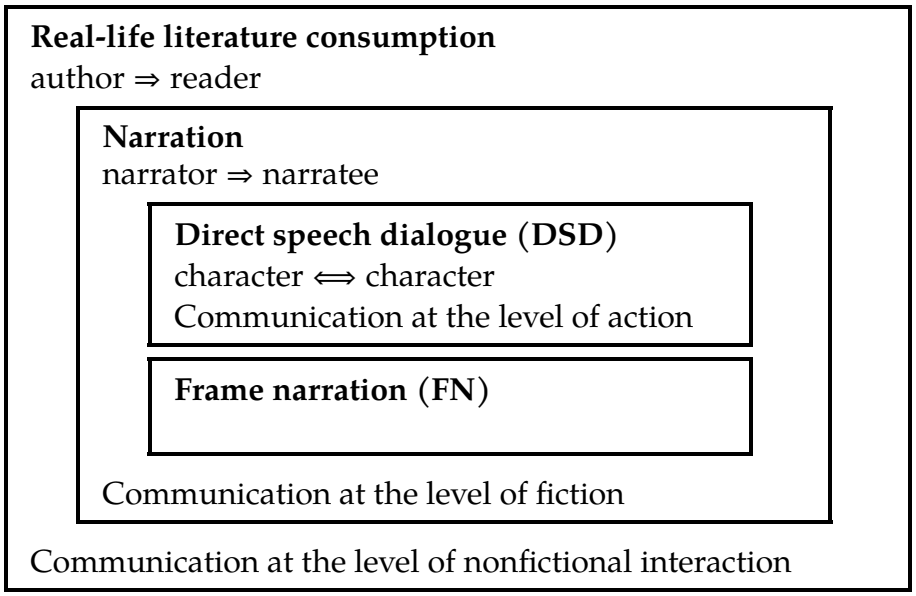

Figure 1: Levels of communication. The box model idea builds on Jahn (2021), which in turn builds on Chatman (1978). Our diagram highlights the communicative status of character-character dialogue.

2.2 The Contrast between Narration and Dialogue

Before we turn to the analysis of dialogue, a few elementary aspects of literary narration need to be sorted out. At a basic level, we have a real-life author, a real-life reader, and a real-life text. At this level, a work of fiction is recognised as such. When being enjoyed by a reader, a work of fiction also introduces a narrator, a narratee, and a narrative. The narrator has the option (as does, by extension, the author) to tell the story by means of reporting what characters have said using the device of direct speech. Or as Genette puts it: "The narrator pretends literally to give the floor to his character" (Genette 1980, 173). Typically, the direct speech dialogue (henceforth: DSD) is surrounded by a 'frame narration' (FN) (Sternberg 1982) of non-dialogue narrative text, which often sets the scene for the dialogue (see Figure 1).

We suggest that there is a prototypical mode of storytelling which is typical for pre-modernist and 'genre' fiction and which is characterised by certain conventions. These can be unpacked as follows:

- There is a balanced interplay between fairly short passages of frame narration (FN) and dialogue.
- The dialogue is mostly rendered as direct speech and follows a style assumed to be close to natural conversation. It has the function to bring a sense of realism to the novels, and to let the characters speak in their own characteristic ways. This introduces stylistic contrasts between the direct speech dialogue (DSD) and the FN as well as among the different characters.
- There is a clear separation between FN and DSD, typically made explicit by the use of quotation marks, dashes, and other linguistic means.
- The dialogue is to a large extent plot-oriented and interactive. Key events play out in conversations. DSD is characterised by a directness, i.e. a "syntactic and deictic

independence of the frame" (Sternberg 1982, 111).

- The dialogue is assumed to be recorded *verbatim*. There is no assumed 'utterance reporter', which can be unreliable, like a narrator.

Examples of literary devices which deviate from these norms are consequently extended monologues, unreliable reporting of conversations, and extended passages of narration without intervening dialogue. The communication among characters can also be captured by other means than direct speech. The two schoolbook alternatives are indirect speech and free indirect speech. Such techniques signal that the reporting of utterance content is done from the narrator's perspective and brings it more in line with other components of the narration, e.g. as regards the narrator's reliability. Speech acts can also be described in ways that only summarise their content or highlight other aspects.

2.3 Defining Genres

Genre is a key parameter in this study. We understand genre as a category based on textual features, authorial intentions, marketing and book-trade practices, and reception (Berglund 2021; Squires 2017). Readers' expectations and authors' and publishers' awareness of readers' expectation play a prominent role for how genres operate in the literary marketplace. In line with Claire Squires, we understand genre as "a crucial component in the marketplace, as it is one of the primary means by which authors and readers communicate" (Squires 2017, 70).

In addition, there is a need to distinguish between popular genres and the genre of prestige literary fiction. As John Cawelti notes, popular genres (or "literary formulas", as he also calls them) "are ways in which specific cultural themes and stereotypes become embodied in more universal story archetypes" (Cawelti 1976, 6). These genres are in contemporary publishing often called simply 'genre fiction'. However, the concept of literary fiction regularly serves as a contrast to popular genre fiction, and it is often deemed "aesthetically interesting, complex and challenging" (Berglund 2021, 760) in other ways than genre fiction. The division between genre fiction and literary fiction thus holds normative assumptions connected to literary quality and accessibility by many agents in the literary ecosystem.

This division is a simplification, naturally, as there are books in between which can be labelled middlebrow or midcult. Even so, the division between the literary and the popular still matters – for authors' careers, for readers, for attention in the media, as well as for the book trade and prize committees. Only by looking at a book cover, it is in most cases easy to distinguish if a book is marketed as 'genre fiction' or 'literary fiction'.

3. Method and Materials

3.1 The Marking of DSD and FN

In Swedish prose fiction, there are two main ways in which DSD is marked as such: firstly, with a quotation mark orthography (as in most contemporary English fiction), or, secondly, with a dash-based orthography. This can be exemplified as follows:

"Too many results", he says. "Time's running out."

– Too many results, he says. Time's running out.

For a computer, just as for a human reader, the interpretation of the dash-based orthography is trickier and requires some work. In novels with such orthography, there is often an unmarked alternation between frame and DSD segments in the same paragraph, which would have been unambiguous if quotation marks had been used. In the second example above, the *he says* clause belongs to the frame narration, but it could also – quite easily if we disregard the wider context – be misinterpreted as a continuation of the DSD segment. The last segment is ambiguous in the inverse way.

Identifying the spans of DSD segments in fiction with a dash-based orthography has to rely on semantic and pragmatic factors. This interpretation consequently constitutes a non-trivial problem for engineering in natural language processing. For the purposes of the present study, we developed a classifier to tell segments of the two kinds, FN and DSD, apart (see subsection 3.3). By contrast, identifying dialogue in quotation mark orthography is with few exceptions a simple task, as the text, so to speak, includes its own explicit markup. Since we are dealing with professionally edited texts, irregularities, such as missing quotation marks, are very rare. The dominating style in contemporary Swedish is to use the quotation mark ” to indicate both the beginning and the end of a segment of DSD. To some extent, this device is also used for quotation of linguistic expressions and to report the content of mental states. However, we can see that such uses are much more infrequent than proper cases of DSD and assume that they only have a marginal and unsystematic impact on the quantitative results reported and analysed here.

3.2 The Corpus

The corpus is intended to form a large and representative sample of fiction that has made an impact in recent years in Sweden by being widely read and/or critically acclaimed. To be more precise, it is a collection of works in Swedish which have been commercially successful in Sweden and/or nominated for the August Prize, the most prestigious award for Swedish fiction. We also included a batch of books written directly for the audiobook medium, but only studied as text data here. Works popular as audiobooks are of particular interest in this context since dialogue is often claimed to be more important in the audiobook format (Åberg 2020; Berglund and Dahllöf 2021).

The corpus includes translations into Swedish, but covers only fiction intended for an adult audience (and thus excludes 'young adult' fiction, children's literature, non-fiction, and non-prose). It comprises 533 books distributed over the following four subcorpora:

- 325 bestsellers in the Swedish book market 2004–2020 based on the annual bestseller charts compiled by the book-trade journal *Svensk Bokhandel*, either on the hardback or paperback top-list. These are dominated by popular genre fiction, especially crime fiction, but also include literary and prize-awarded fiction.
- 56 beststreamers in the Swedish book market 2015–2020, i.e. novels which have been among the twenty most streamed audiobooks each year in the leading online book streaming platform Storytel, but that were not bestsellers in print.

- 54 titles shortlisted for the August Prize (including winners) 2004–2020 that were not bestsellers in print. (This subcorpus includes all shortlisted titles but four, for which no e-book version has been published.)
- 98 'born-audio' works, viz. all adult fiction in Swedish published in the 'Storytel Original' series from its inception in 2016 until May 2021. (These works are written for online audio streaming, but are also available as e-books.)

We analyse our data in terms of four categories: crime fiction, romance, prestige fiction, and other fiction. Two of these – crime fiction and romance – belong to well-established literary genres, while the two remaining are categories of fiction of a different kind. We apply a literary sociological approach and rely on genre categorizations and classifications made by key agents in the literary field (publishers, literary academies, and prize committees) to operationalise genre labels and literary prestige that align with our theoretical understanding, as follows:

The categorisation of 'crime fiction' follows the bibliography of Swedish crime fiction compiled by the Swedish Crime Fiction Academy, which applies an inclusive genre definition in line with book trade categorisations. This is a thematic definition that covers all kinds of crime related plot structures and subgenres, including detective stories, police novels, and psychological thrillers (Berglund 2012, 40–41).

We use the label 'romance' for an array of popular genre fiction that includes romance (in the more restrictive sense) and all its sub-genres, chick lit, and women's fiction (which in Sweden is often labelled 'feelgood'), complying with, for instance, the Romance Writers of America (2025).[2]

As 'prestige fiction' we count all titles that have received or been nominated for (and possibly won) a major literary award, which has been operationalised as the Nobel Prize in literature, the Pulitzer Prize, the Booker Prize, and the two most important literary prizes in a Swedish context: *Augustpriset* and *Nordiska rådets litteraturpris*. It should be noted that this is an operationalisation of literary prestige that is likely to exclude some titles that probably would be regarded as prestige or literary fiction by some critics.

The label 'other fiction' applies to all titles which do not belong to any of the three specific categories. It thus covers popular genre fiction other than crime fiction and romance (e.g., fantasy, science fiction), middlebrow literary fiction as well as any literary work overlooked by the relevant prize committees.

The composition of the corpus in terms of genre and subset (by selection criterion) is shown in Table 1, while Table 2 gives an overview in terms of genre and DSD-related orthography type. As becomes apparent, our corpus is not well-balanced in terms of literary categories and genres, most notably since crime fiction is so dominant. However, this reflects the dominance of crime fiction on the bestseller lists in Sweden. In a literary-sociological sense, then, if one is interested in the kinds of fictions that most people read in Sweden, our corpus presents a fairly well-balanced picture.

2. There is no bibliography that singles out romance in this wider sense in Sweden. Rather, the categorisation has been made manually for all titles that by their publishers have been labelled with a Thema code markup of romance of any kind ('FR' – fiction → romance – and its sub-categories, Editeur 2025), and/or that have peritextual attributes clearly connected to the tropes of romance (i.e. book cover texts including keywords like *romance*, *love*, *feelgood*, *heart-warming*, *dating*, and/or book cover images and design showing genre-typical combinations of couples or females, quirky handwriting and bright pastel colours).

	Bestseller	Original	Beststreamer	August	Total
Crime fiction	187	43	53	0	283
Romance	35	30	3	0	68
Prestige fiction	32	0	0	54	86
Other fiction	71	25	0	0	96
Total	325	98	56	54	533

Table 1: Composition of the corpus regarding category/genre and subset: number of books.

	Quotation marks	Dash	Other	Total
Crime fiction	184	99	0	283
Romance	61	7	0	68
Prestige fiction	39	24	23	86
Other fiction	68	20	8	96
Total	352	150	31	533

Table 2: Composition of corpus regarding genre and DSD-related orthography type.

Of the 533 books in the corpus, 32 books are excluded from the full quantitative analysis, in most cases because they do not contain substantial amounts of clearly marked DSD. In 31 of these cases, the books do not contain substantial amounts of DSD marked in a clear or conventional way. Carola Hansson's *Minnestrådar* (2019) represents another kind of exception. Here, the DSD form is used as an overarching narrative technique, where one character (Madame Vidal) is retelling the main story as it appeared from her memory to two listeners, who are only minor characters in the narrative as a whole. Since the narrative elements are more or less completely embedded in the DSD in Hansson's book, it becomes an extreme outlier for all measures concerning direct speech, which makes it less productive to analyse quantitatively (although the monologue form of Madame Vidal's retellings of course ontologically has the status of a dialogue). For this reason, i.e. its extreme outlier position and its use of DSD which is unlike all other novels in the corpus, it has been excluded from the quantitative part of the study, for which 501 books are left. Notably, the 32 titles excluded from the quantitative analysis all belong either to the category of prestige fiction (24) or to other fiction (8).[3]

3.3 Separating DSD and FN Segments

As mentioned above, we used an automatic classifier to separate DSD and FN text in novels with a dash-based orthography. For works with an orthography based on quotation marks, this separation is in general a trivial matter, as quotation marks in principle provide an explicit markup. This means that professionally edited native quotation mark text can easily be converted into correctly annotated dash-style data (see subsection 3.1 for an example). We can use such minimally artificial data for supervised training and validation of classifiers working on documents in dash-based orthography (Jannidis et al. 2018; Kurfalı and Wirén 2020). The development of the classifier which we used to separate FN and DSD text in paragraphs originally written in dash-based orthography relied on training data of this kind.

3. A listing of the specific reasons for excluding each of these 32 books from the computational analysis is found in the Supplementary Materials, in the file gddcorpex.csv.

The exploration and training of a classifier for the dash-based orthography followed the procedure of Dahllöf (2022). Accordingly, we distinguished the following three categories of paragraph:

- Typical DSD paragraphs, i.e. those beginning with DSD material as indicated by an initial quotation mark (") or dash (–). They are often DSD only, but alternation between DSD and FN text is also common. Henceforth: DSD/FN paragraphs (The classifiers we trained targeted the DSD/FN separation in the dash-orthography variety of such paragraphs).
- FN only paragraphs, i.e. those not beginning with a dash or containing quotation marks. Henceforth: FN paragraphs.
- Other paragraphs, i.e. those with non-initial explicitly quoted text or non-matching quotation marks. They were not included in the training and application of classifiers.

The machine learning design departed from textual data tagged with part-of-speech (POS) labels. We used the Stanza (Peng et al. 2020) tagger (version 1.4.2) with the "talbanken 1.0.0" (Nivre and Smith 2008) model for Swedish to obtain the POS labelling. We did not explore the machine learning possibilities in a comprehensive way, but we found a classifier design which we considered satisfactory for the purposes at hand. The approach was based on classifying snippets of sentences using a multilayer perceptron (Dahllöf 2022). The paragraphs were segmented at punctuation marks, as these are the only points where DSD/FN shifts may occur under a conventional orthography. Features were derived from both the snippet text and the preceding narration context. They involved POS unigrams and bigrams, word forms, as well as pronoun person and verb tense. This makes it possible for the classifier to be sensitive to both lexical information and point of view differences between the DSD and the FN. We adapted the data size decisions to the current corpus and applied a cross-validation setup (as explained in Dahllöf (2022)) which was three-fold over 117 + 117 + 117 = 351 books (compare Table 2) taking, for training, 90,000 randomly selected snippets (each one labelled as FN or DSD) from 234 books and, as testing data, all snippets (both FN and DSD ones) from all dialogue paragraphs from the 117 other books. We tried a number of multilayer perceptron network designs, in terms of number of layers and number of neurons in each layer. One with three layers, comprising 13, and 11, and 9 neurons, respectively, was found to be the optimally performing architecture among those considered in Dahllöf (2022), with 800 features selected. This classifier yielded accuracy scores from 94.1% to 94.2% over the folds (on word token level: 92.9% to 93.2%). (This is slightly lower than in Dahllöf 2022.) We used the best-performing classifier in terms of accuracy to tag the novels in dash-based orthography. We evaluated its performance by manually checking selected snippets from the 150 relevant novels in our corpus. One in 2,500 snippets were randomly added to the sample. This yielded an accuracy score of 96.5%, 301 snippets of 312 being correctly labelled. On the token (word instance) level, the accuracy score was 94.4%. (Both scores are slightly better than those recorded in Dahllöf 2022.) In Table 3, we report the precision and recall scores.

The results discussed below are based on book-level measurements of the relative

	Precision (snippets)	Precision (words)	Recall (snippets)	Recall (words)
DSD	99.6%	99.6%	96.1%	93.8%
FN	85.1%	73.1%	98.3%	97.8%

Table 3: Precision and recall scores.

Genre	DSD share	DSD length
Prestige fiction	14%	10.7
Crime fiction	30%	12.6
Romance	31%	10.5
Other fiction	23%	12.7
All	27%	12.1

Table 4: Amount of DSD and word length of DSD insets: mean per genre.

amount of dialogue and the mean length of the dialogue insets (in words), the share of dialogue for each segment when the text is divided into ten equally long parts (based on word count), and the corresponding values when text is divided into twenty equally long parts.[4]

4. Results and Discussion

4.1 Amount of Dialogue

As expected, genre is a key parameter reflected in the relative amount of DSD, as calculated per word, i.e. all words in all DSD insets divided by the book's total word count. In particular, there is a clear tendency that prestige fiction contains less DSD on average (14%) than crime fiction (30%) and romance (31%). The category "Other fiction" lies somewhere in between (23%) (Table 4).

That DSD is a more important component in popular genre fiction than in more aesthetically oriented and experimental literary fiction aligns with previous research (Berglund et al. 2019; Cawelti 1976; Gelder 2004), but the extent to which DSD appears as a more important feature of popular genre fiction is surprising. There are no romance or crime novels with a DSD share below 10 percent – it ranges from just over 10 percent to approaching 60 percent. For the category of prestige fiction, on the other hand, 23 of the 62 books (37%) have a DSD share below 10 percent, and 52 (84%) have a share below 20 percent. Prestige fiction with higher amounts of DSD – say, over 30 percent – is uncommon, there are only 4 such novels in the subset (7%), and only one of them has a DSD share of over a third. Again, this can be contrasted to crime fiction and romance, where 92 (33%) and 23 (34%) of the novels in the respective genre consist of at least a third of DSD, and there are plenty of examples where it approaches or even passes half of the books' total word count.

Thus, we find a negative correlation between the amount of DSD in prose fiction books, and the status they are given in the literary field: the lower the amount of DSD, the more likely a title belongs to the category of literary fiction.

4. All of these values for every book are found in a file among the Supplementary Materials.

As we saw above, the 32 books which were excluded from the quantitative analysis because they do not contain substantial amounts of clearly or conventionally marked DSD also conform to this inverse relationship, as they are works of prestige fiction (24) or other fiction (8).[5]

These general observations need to be nuanced. The books of prestige fiction with above 30 percent dialogue appear to rely on a kind of storytelling that uses DSD in a straightforward way, similar to how it is used in most popular fiction. Notably, three of these four outliers of prestige fiction are historical novels – Carl-Michael Edenborg's adventure novel *Alkemistens dotter* (*The Alchemist's Daughter*) (2014; 34% DSD) set in the late 18th century; Anna Fock's novel *Väderfenomen* (*Weather Phenomenon*) (2019; 31%) about how a nuclear plant disaster in Soviet Ural in 1957 affects generations to come; and Nobel prize laureate Kazuo Ishiguro's *The Remains of the Day* (originally published 1989; 31%), retelling the memories of a butler loyally serving in a near-Oxford manor in the 1920s and 1930s. The fourth outlier is Donna Tartt's *The Goldfinch* (2013), written by an author famous for epic and wide-spanning novels.

In the reception of these novels in Sweden, their reliance on narrative techniques similar to those used in popular genre fiction is often explicitly discussed. For instance, Fock's novel is said "to borrow effective techniques from suspense fiction" and "to get energy and nerve by means of conventional literary techniques" (Folkhammar 2019), Edenborg is characterised as writing "suspenseful as Dan Brown" (Kärnborg 2014), while Tartt is claimed to have written "a thriller" (Högström 2013) that "excels in long, precisely captured dialogues" (Malmberg 2013).[6]

The reason for these novels' prestige in the literary field, then, have little to do with experimenting with literary form, and more with their content and the quality of their straightforward storytelling. Prestige fiction, then, cannot be said to be characterised by a low amount of DSD generally, since there are several counterexamples to this.

By contrast, popular genre fiction generally has high amounts of DSD. The opening scene in Denise Rudberg's *Under nio nätter* (*During Nine Nights*, 2019), the book with the highest amount of DSD in the corpus (55.9%) can be used as an illustration of this way of writing popular genre fiction:

> Torsten stepped shakily through the gate and jumped when a neighbor rushed in behind him before the door slammed shut. The neighbor, a lady in her seventies, smiled apologetically.
>
> – Sorry, I didn't mean to scare you. But maybe you're not that easily scared, because surely you're the police?
>
> – Yes that is correct. I moved in a while ago.
>
> The lady lowered her voice:
>
> – I have heard that you have some kind of secret agent activity here in the house.

5. See the Supplementary Materials for details on each book.
6. The quotes from the Swedish reception have been translated into English by us.

> – Well, it's not really that exciting. We mostly do archiving of old reports. A kind of side business so that the real police can do their jobs, so to speak. We are the ones left over when the police were restructured. But not against me, I enjoy sitting outside, so much nicer premises. Not to mention the coffee – what a difference. (Rudberg 2019, 13–14, our translation)

Popular genre fiction with smaller amounts of DSD works differently. In the bottom segment, below 20% DSD, we find books by a blend of romance and crime writers. When looking at those books we find a particular variety of storytelling in a number of cases. Consider the opening scene of Sofie Sarenbrant's *Bakom din rygg* (*Behind Your Back*, 2017), one of the popular genre fiction books in the selection with the smallest amount of DSD (11.3%):

> The shame of being late turns to irritation the moment she steps through the door of the legendary Salon De Luca on Riddargatan.
>
> Her client is sitting in the wrong chair.
>
> Thirty-one-year-old Jenny Ericson understands why Mia Bonnevier desperately reaches for her cell phone as soon as they see each other. She is of course embarrassed.
>
> "Well, hello," Jenny says questioningly and adjusts her patterned blouse from Versace. Mia greets back nervously and averts her gaze.
>
> You traitor.
>
> Of course she can have her hair cut by whoever she wants. Customers come and go, even at Stockholm's hottest hair salon, and it's important not to take it too personally if someone wants to switch. Still, Jenny sees it as a defeat, especially when they prefer Angelina Silver. It would have made a difference if they chose a different salon, but why would anyone? She looks around the beautifully decorated room with high ceilings, well-preserved parquet floors, gold frames around the mirrors, black and white portraits on the walls and generous space for just three chairs. This has been her workplace for three years now, her second home. But it doesn't feel as fun to come here anymore. (Sarenbrant 2017, 11–12, our translation)

In comparison with Rudberg's scene, four things stand out as distinctly different:

1. Where Rudberg has rather extensive DSD insets often spanning over several sentences, Sarenbrant has short ones, involving at most one clause.
2. Where Rudberg uses external focalisation (Genette 1980, 189–194), Sarenbrant instead uses internal focalisation, in the example the story is focalised through Jenny's point of view.
3. Where Rudberg writes DSD in a straightforward way without engaging in any major description of how things are uttered, Sarenbrant on the contrary repeatedly filters what is said through Jenny's subjective perspective, often using elaborate reporting clauses and so-called suspended narration.

4. Where Rudberg has little or nothing in the way of interior monologue or subjectively filtered frame narration, such narrative techniques are prominent in Sarenbrant's book, both in direct relation to DSD, and in longer stand-alone paragraphs.

Thus, although both Rudberg's and Sarenbrant's novels have many things in common on the surface level – they are commercially successful and plot-oriented third-person narrative crime fiction books, written by female Swedish authors and published in the same period of time – their narrative techniques are fundamentally different. Rudberg uses external focalisation and lots of DSD, whereas Sarenbrant uses internal focalisation and less DSD. Other books with low DSD rates, such as Håkan Nesser's *Skuggorna och regnet* (*The Shadows and the Rain*, 2004) and Malin Persson Giolito's *Störst av allt* (*Quicksand*, 2016) use internal focalisation in much the same way as Sarenbrant, although these books are first-person narratives. This suggests that the correlation between the amount of DSD and the overarching kind of focalisation technique deployed might hold true also on a more general level in popular genre fiction.

These observations connect to the argument made by Wilkins et al., who contrast the "sharp fast-paced dialogue" of a crime fiction novel with the "introspective" and "subjective" mode of storytelling of romance (Wilkins et al. 2022, 163–164). This contrast is in many ways what we see in the case of Rudberg and Sarenbrant. But while Wilkins et al. argue that these narrative techniques are firmly rooted in the respective genre, our data indicate something else. In all genres included in our corpus, it is most common to write in a reporting and journalistic way, which equals external focalisation and high levels of DSD, making subjective character focalisation less important. But it is also, regardless of genre, perfectly possible to write the frame narration in a subjective and introspective way with internal focalisation, which often leads to a smaller share of DSD. Internal focalisation also gives the author the option of using an unreliable narrator, as, for instance, does Gillian Flynn in *Gone Girl* (2012) and Paula Hawkins in *The Girl on the Train* (2015), both included in the corpus.

4.2 The Distribution of Dialogue

On average, DSD is rather evenly distributed over the books in the corpus. If we divide each book in the corpus into twenty equally long parts (based on word count), we can note an almost flat line, with a moderate increase of the share of DSD over the text.[7] Exceptions to this are that the introductory and concluding parts have smaller fractions of DSD (see Figure 2).

The tendency to lower shares of DSD in introductions makes intuitive sense. Although novels can start *in medias res*, it is common to dedicate the narrative beginning for setting the scene or for explaining the background to the narrative to come. In crime fiction, the first chapters are often used to introduce the mystery by telling the narrative from the point of view of the murderer, the victim, or someone who discovers a dead body. Mari Jungstedt, for instance, uses this technique frequently. In *Den dubbla tystnaden* (*The*

7. The exact values for each book are found among the Supplementary Materials. However, it should be noted that the lengths of the books in our corpus vary considerably – from the shortest book containing 14,235 tokens to the longest containing 378,489 tokens. This means that each of the 20 parts of the novels used to calculate distribution also vary in length in absolute numbers between the novels accordingly.

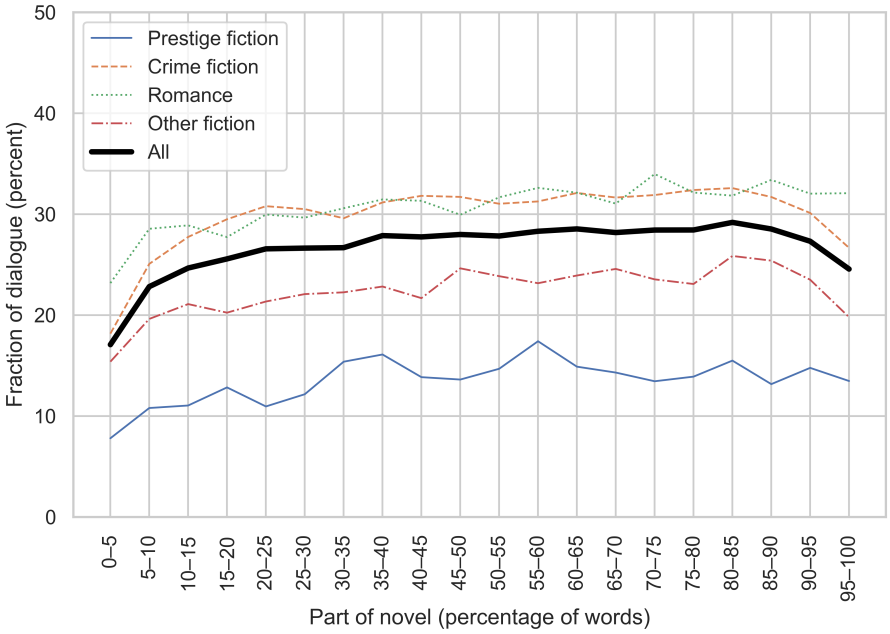

Figure 2: Distribution of DSD per novel twentieth and genre.

Double Silence, 2009), Jungstedt uses the beginning of the book to set the scene for the remaining narrative by focusing on the killer and the victim in the first chapter, and on two characters around which the murder plot is spun in the following two chapters. These three chapters have almost no DSD at all, and serve as a background for the rest of the book, where there is much more DSD. Similar kinds of introductions with a low share of DSD are found in many crime novels in the corpus.

Other genres can exhibit the same pattern. In Chris Cleave's *The Other Hand*, the first 5,400 words are spent on setting the scene and introducing the protagonist and her background for the reader. Then, the plot starts and with it a significant DSD increase (Cleave 2008, 7–9).[8]

In Jenny Colgan's *Christmas on the Island* (Colgan 2018), the first two chapters are spent entirely on describing the landscape on the small Scottish island where the novel is set.[9] Novel beginnings, then, appear to have a rather special status in relation to DSD and narrative techniques used.

There is also a tendency to lower shares of DSD in the final parts. This indicates that dramatic finales appear to be formally different from the main narrative. Interestingly, this holds true only for crime fiction and 'other fiction', and not for prestige fiction and romance (see Figure 2).

An explanation can be derived from the two main drivers of reader interest within crime fiction: *curiosity* (who did it?) and *suspense* (what will happen?) (Todorov 1977). In contemporary crime fiction, the murderer generally gets exposed towards the end of the narrative, and at that point the curiosity-driven narration is transformed into a suspense-driven one (Berglund 2017, 48). This is reflected in Figure 2. The investigation

8. The first tenth's DSD share of 7.2% can be compared to the 32% mean for the entire novel.
9. The 6.5% DSD share of the first tenth in the novel can be compared to the 32% DSD share of the novel as a whole.

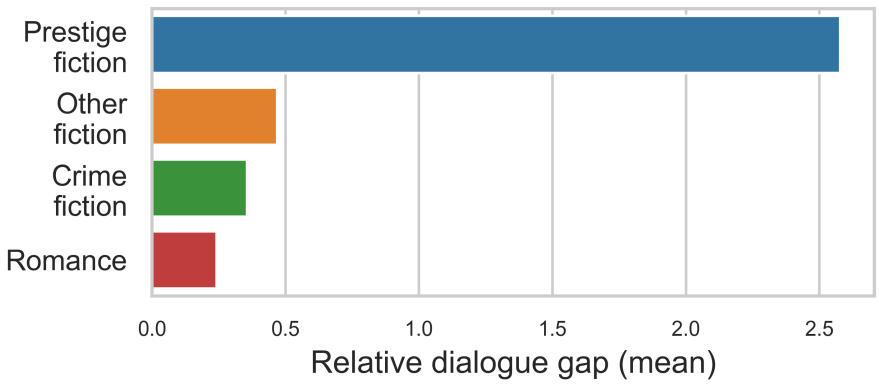

Figure 3: Average relative dialogue gap per genre.

ends and the narrative turns into a hunt for the killer or a car chase or some other kind of explicitly thriller-oriented plot, where dialogue is less apparent than in the narrative as a whole. Romance works differently. Here, the final plot drama or twist rely mostly on the relational level between the characters, and less on action-based events, which keep the high DSD level intact to the end.

Text averages do not capture differences within individual works. To understand what is going on as narratives unfold, we need to analyse the distributional variation of the DSD fraction over the texts. One way of quantifying such variation is to measure the ratio between the tenths in the narrative with the highest and lowest fractions of DSD. This gives us a metric of the 'relative dialogue gap' within narratives. Although a part of this gap is explained by differences present in many narratives – where beginnings and endings have lower amounts of DSD than other parts, as we show in Figure 2 – what the method robustly captures are extreme variations in DSD distribution.

What immediately stands out is that prestige fiction has higher levels of intratextual variation in the DSD share than the other genres. On average, prestige fiction shows a gap of 2.6 percentage points, whereas the corresponding scores for all other genres are below 0.5 (0.58 for other fiction, 0.36 for crime fiction, and 0.24 for romance) (see Figure 3).

This kind of variation is thus more pronounced in literary fiction. When looking at the outliers for this metric, all of them are literary fiction, and all interleave long parts lacking dialogue with long parts where dialogue is prominent. In Per Gunnar Evander's partly autobiographical novel *I min ungdom speglade jag mig ofta* (*In My Youth I Often Looked in the Mirror*, 2005) – the most extreme outlier in the corpus with a relative dialogue gap of 46% – the first 30% segment has no dialogue at all. The novel is told in the first person and memories, thoughts and everyday whereabouts are internally focalised through the protagonist/narrator, a middle-age author resembling Evander. The first direct speech instance comes on page 83. It is the protagonist's neighbour speaking, as reported through the memory of the protagonist. After this point, however, passages of DSD are repeatedly interleaved with internally focalised contemplating narration. Importantly, the dialogue is not used for driving the plot forward – it is in a sense hard to talk about a formal plot in this novel – but rather as a stylistic narrative feature for addressing certain topics. For instance, in the middle of the novel, the protagonist and

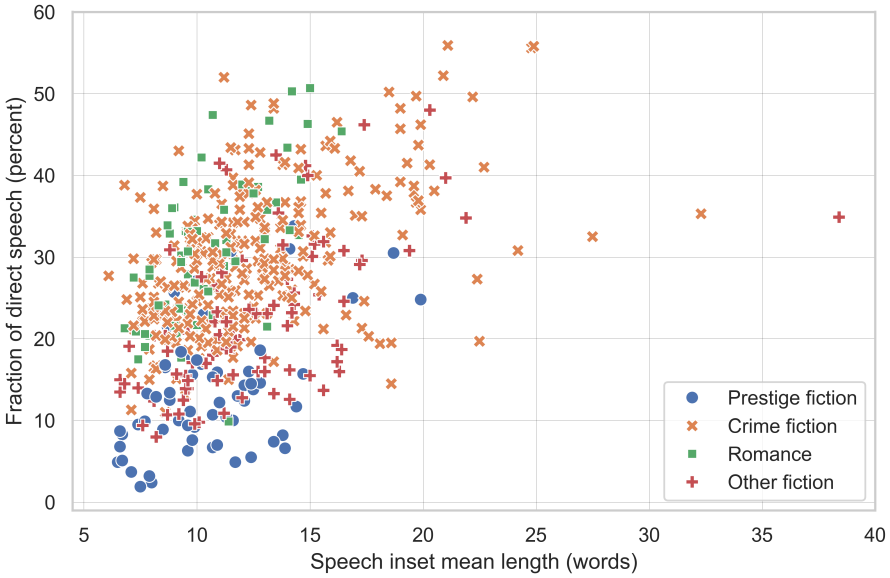

Figure 4: Books with genre indicated: mean length of DSD insets by fraction of DSD (n=501).

the neighbour Eskil have a long conversation on a fishing trip regarding the truth claim in the protagonist's novels. Here, DSD is a way for Evander to address the question of fact versus fiction, which can be said to be a theme in the novel. The realistic feature of the DSD is crucial as Eskil has the function of speaking as a commoner, as opposed to the novel's I, who is an author.

4.3 The Lengths of Dialogue Insets

In Figure 4, the books in the corpus are plotted based on the share of DSD and the average length of DSD insets (cf. Table 4).

At least three tendencies can be discerned in Figure 4:

1. There is a considerable variation in the mean length of DSD insets per work (spanning from 6 to 39 words). This indicates that authors use DSD in very different ways.

2. While there is a positive correlation between amount of DSD and length of DSD inset, this correlation is weak to moderate (in statistical terms, the Pearson correlation coefficient is 0.45).

3. Genre appears to be an important factor also when it comes to DSD inset length. Most notably, romance books have somewhat shorter insets on average (10.5 words) than crime fiction (12.6), and all of the outliers with very long DSD insets are either crime fiction or works in the other fiction category. Even if romance is the genre with the highest amount of DSD, it has the shortest DSD insets, even shorter than prestige fiction (10.7) (see Table 4).

Mean length of DSD insets also seems to capture individual traits in how authors handle dialogue. Most authors in the corpus show small variation both in the average share of DSD and in average length of DSD insets. Dan Brown, for instance, with six books in the corpus, exhibits a variation in mean length of DSD insets from 10.7 words (*Angels*

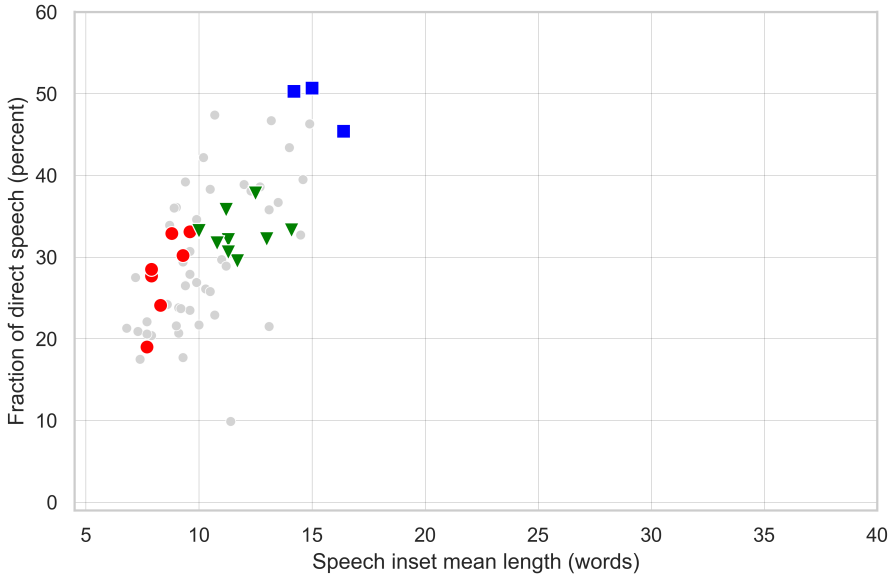

Figure 5: Mean length of DSD insets by fraction of DSD for romance, with the books of three romance authors highlighted: Jenny Colgan (red dots), Jojo Moyes (green triangles), and Lucinda Riley (blue squares).

and Demons) to 13.0 (*Origin*). The variation among Lars Kepler's nine thrillers in the corpus is even smaller, spanning from 7.2 words (*The Hypnotist*) to 9.0 (*The Mirror Man*). Conn Iggulden's two historical novels have almost identical DSD mean lengths: 16.2 words (*Bones of the Hills*) and 16.5 (*The Gods of War*). The same goes for Gillian Flynn's thrillers *Dark Places* (14.8) and *Gone Girl* (14.9) and for Kazuo Ishiguro's two literary novels *The Remains of the Day* (18.7) and *Never Let Me Go* (19.9). We see that authors often have their own particular way of writing dialogue, and they appear to stick to it across books. This holds true for most authors in the corpus, for popular genre fiction as well as literary fiction, and for those who write a lot of dialogue as well as for those who use it more sparsely.

Authors appear to simultaneously follow genre patterns and their own personal practices. This shows in Figure 5, Figure 6, and Figure 7, where the markers highlighted in red, green, and blue are books from specific authors (authors' own practices), and the grey dots are all other books in the same genre or category (the larger genre pattern). The romance genre, visualised in Figure 5, in general has lots of dialogue but rather short insets, but within the genre there is a spread in DSD use from short insets (e.g. Jenny Colgan) to longer insets and more dialogue (e.g. Jojo Moyes) to long insets and high levels of dialogue (e.g. Lucinda Riley). A similar spread can be seen in crime fiction, visualised in Figure 6, from short insets (e.g. Lars Kepler) to mid-length (e.g. Mari Jungstedt) to very long (Niklas Natt och Dag) as well as in literary fiction, visualised in Figure 7, from short insets (e.g. Per Olov Enquist) to mid-length (e.g. Elena Ferrante) to long (Kazuo Ishiguro).

Short DSD insets are often a means for writing rapid dialogue sequences as a way of moving the plot forward in popular genre fiction. But there are exceptions. In Niklas Natt och Dag's historical crime novel *1793* (2017), lengthy utterances are used to explain both plot-related and historical events. Here is an example:

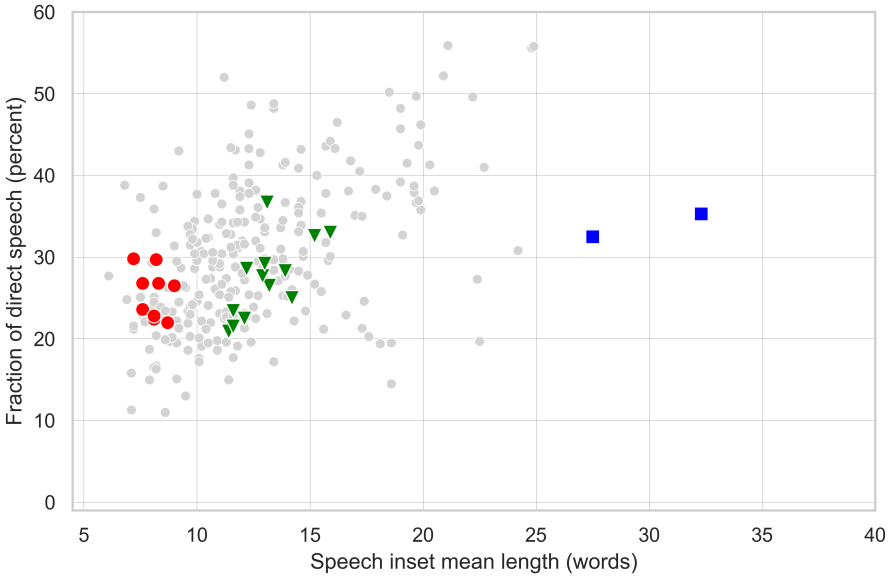

Figure 6: As Figure 5 but for crime fiction, with three crime authors highlighted: Lars Kepler (red dots), Mari Jungstedt (green triangles), and Niklas Natt och Dag (blue squares).

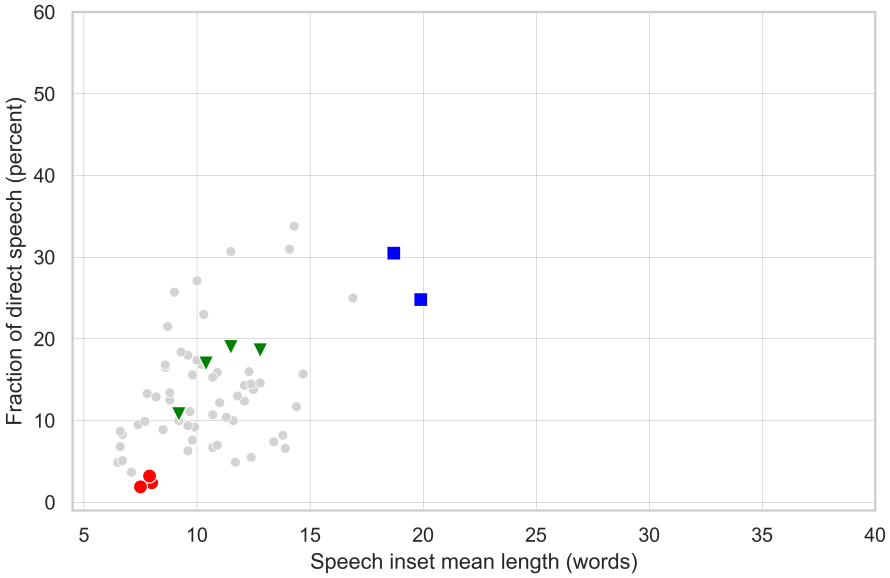

Figure 7: As Figure 5 but for literary fiction, with three authors of literary fiction highlighted: Per Olov Enquist (red dots), Elena Ferrante (green triangles), and Kazuo Ishiguro (blue squares).

> Norlin holds up a folded paper with a freshly broken seal.
>
> "This is a letter signed by Gustaf Adolf Reuterholm, in which he demands to know why no progress has been made in the investigation he demanded regarding a rumor that he has tried to poison the crown prince. The same rumor claims that his hunger for power can be traced back to impotence and a long list of perverse tendencies. The baron feels that he has waited long enough to see those responsible taste the rod, and now demands that I provide him with a full account of my efforts."
>
> "And will you send him one?"
>
> "Since I haven't done anything, it would probably be best not to. The man is out of his mind. Reuterholm is nothing but a despot, without friends or family to provide a sense of stability. He's trying to get the fortune-teller Arvidsson to speak with the dead on his behalf. Vain, testy, and resentful to boot, just like King Gustav himself as time went on. Fear of revolution and betrayal is a pestilence that spreads to all whose posteriors come too close to the throne. [...] While King Gustav had nightmares of the revolution in France spreading to the far north and did everything in his power to eavesdrop on republican chatter in the coffeehouses, his killers were sneaking around among members of his own court. He was so afraid of the commoners he never met that he believed the nobles – right in front of his face – were harmless." (Natt och Dag 2017, 45–46)

This kind of dialogue lacks realism in King's sense – the mini-lecture on Swedish history seems more directed to a 21st-century reader than to an 18th-century person, and long-winded subordinate clauses hardly give the impression of spontaneous conversation. Still, Natt och Dag's dialogue has some kind of 'realistic quality' since his long DSD insets sound old-fashioned and thus 'historical' to a 21st-century reader. It somewhat imitates 19th-century novels in a way that potentially is a character trait of historical fiction more broadly and that would be interesting to investigate further.

5. Conclusions

In this article, we have investigated direct speech dialogue (DSD), the 'audio portion' of novels, according to King (2020). Although most dialogue indeed clings to the norm that it in some respects should reflect how people talk, in Stephen King's sense, there is considerable variation in how, to which extent, and for what purposes authors make use of dialogue. We approached dialogue quantitatively through a study of 500 books of contemporary prose fiction in Swedish. By focusing on three specific metrics – DSD share, DSD distribution over books, and length of DSD insets –, we were able to pin down several key parameters.

Dialogue is clearly a more commonly used device in popular genre fiction than in literary fiction. Crime fiction and romance contain twice as much dialogue as literary fiction on average, even if there are exceptions to this pattern. Similarly, it is clear that literary fiction is more varied when it comes to how the dialogue is distributed over the stories. In popular genre fiction, the proportion of dialogue is most often more or less the same

Genre	Share	Variation	Length
literary/popular genre	high/low	high/low	long/short

Table 5: Four binary choices for analysing fiction.

throughout the books, the most common exception being that the beginning contains less dialogue. In literary fiction, dialogue can either be used by authors in a similar and conventional way, or as a distinct aesthetic choice in shorter prose passages while completely absent in others. It is plausible that authors of literary fiction make use of indirect speech and free indirect speech to a larger extent than popular genre fiction. Thus, the main difference might not be between if characters speak or not, but rather how their speech is reported by the author.

The results of our investigation did not agree with the assumption in previous research that crime fiction is more dialogue-driven than romance. On the contrary, books categorised as romance contain more dialogue on average than those categorised as crime fiction, although the differences were small.

The third parameter we investigated, average length of dialogue insets, also relates to genre, but to a lesser degree. Instead, what seemed clear was that authors tend to use dialogue in their own author-specific ways. A vast majority of the books by the 87 authors with more than one book in the corpus clustered closely together by author in terms of average inset lengths (as well as regarding how much dialogue they use in their books). This was perhaps not unexpected, but the strength of this pattern was surprising – not least since it holds true across all genres, including literary fiction. Most books have short to semi-short insets (the average length is 12.1 words), but some authors tend to expand their characters' direct speech utterances into lengthy mini-lectures that appear to be directed more towards the reader than the character to which they are spoken.

The strength of our approach, we argue, lies in its scalability for looking at narratives. If our three measures are simplified into binary distinctions 'high' or 'low', they form a framework that can be applied for describing fiction. With this approach, a novel can be associated with one of eight possible combinations. Since genre is a fundamental factor here, we propose that the distinction between popular genre fiction and literary fiction should be integrated into the framework. This gives us a total of 16 theoretically possible complex categories, as shown in Table 5.

For instance, the default contemporary crime fiction or romance novel has a high amount of dialogue, with little variation in its distribution, and with rather short insets. According to our schema, then, such a book is categorised as high-low-short. One example of this kind highlighted in our previous discussion was Lars Kepler's crime novel *Hypnotisören* (*The Hypnotist*). (Actually, all nine crime novels by Lars Kepler in our corpus share this trait.) But popular genre fiction can likewise contain lots of dialogue and long insets (high-low-long), like crime novels by Niklas Natt och Dag and romance by Lucinda Riley. Books can also be more sparse in their dialogue (low-low-short), as the example with the internally focalised *Bakom din rygg* (*Behind Your Back*) by Sofie Sarenbrant shows. The only positions in the schema we have not encountered in our corpus are popular genre fiction with high variations in dialogue distribution.

Literary fiction can, just like popular genre fiction, have high shares of dialogue, with little variation in distribution, and with rather short insets. One example with such a high-low-short categorisation is Donna Tartt's *The Goldfinch*. But literary fiction can also have lengthy insets (high-low-long) as Kazuo Ishiguro's books, or very sparse dialogue (low-low-short) as Per Olov Enquist's books, or very varied dialogue in terms of its distribution (low-high-short) as Per Gunnar Evander's *I min ungdom speglade jag mig ofta* (*In My Youth I Often Looked in the Mirror*). Again, authors of literary fiction can use dialogue in a more flexible way than authors of popular genre fiction, who cannot stray too far away from readers' expectations. Consequently, literary fiction more frequently employs combinations of direct speech, indirect speech and free indirect speech with higher degrees of originality, and thus establishes modes of storytelling which might be more surprising to the readers.

We believe that the kind of "reductive reading" (Allison 2018) proposed by our framework can enable us to see the larger picture, to identify outliers, and to place novels in a wider context of dialogue handling. The results obtained by such a quantitative analysis can perfectly well be combined with close readings of how the dialogue works on deeper levels. Our hope is that this article has shown a way forward for the study of dialogue on a larger scale, as well as how to connect quantitative methods with more traditional modes of reading.

6. Data Availability

We provide a set of Supplementary Materials. The actual texts of the corpus behind the study cannot be shared due to the fact that they consist of recently published fiction protected by copyright. Instead, we share the corpus data as POS and DSD tagged sequences, one for each book, with the specific graph words removed. The Supplementary Materials also include metadata for the corpus and measurements per book. These data are all stored as comma-separated values (CSV) files. We also provide the Python code for compiling the measurements discussed in this article from the corpus, as well as the code used to generate the figures from these measurements. The Supplementary Materials can be found, along with a more detailed documentation at this DOI: `https://doi.org/10.5281/zenodo.22085344`.

7. Author Contributions

Karl Berglund: Funding acquisition, Conceptualization, Data curation, Methodology, Visualization, Writing

Mats Dahllöf: Conceptualization, Data curation, Methodology, Formal analysis, Investigation, Writing

References

Åberg, Daniel (2020). "At skrive til øret: Betragtninger fra en Storytel Originals-forfatter". In: *Passage* 83, 74–77. 10.7146/pas.v35i83.121609.

Allison, Sarah (2018). *Reductive Reading: A Syntax of Victorian Moralizing*. Johns Hopkins University Press. 10.1353/book.58979.

Allison, Sarah, Heuser Ryan, Matthew Jockers, Franco Moretti, and Michael Witmore (2011). *Quantitative Formalism: An Experiment*. Stanford Literary Lab Pamphlet 1. https://litlab.stanford.edu/LiteraryLabPamphlet1.pdf (visited on 06/08/2026).

Bal, Mieke (2017). *Narratology: Introduction to the Theory of Narrative*. 4th ed. University of Toronto Press.

Berglund, Karl (2012). "Detectives in the Literary Market: Statistical Perspectives on the Boom in Swedish Crime Fiction". In: *Scandinavica: An International Journal of Scandinavian Studies* 51 (2), 38–57. 10.54432/scand/VIUH2465.

— (2017). "Killer Plotting: Typologisk intriganalys utifrån fjärrläsningar av samtida svenska kriminalromaner". In: *Tidskrift för litteraturvetenskap* 47 (3–4), 41–68. 10.54797/tfl.v47i3-4.7852.

— (2021). "Genres at Work: A Holistic Approach to Genres in Publishing". In: *European Journal of Cultural Studies* 24 (3), 757–776. 10.1177/13675494211006097.

Berglund, Karl and Mats Dahllöf (2021). "Audiobook Stylistics: Comparing Print and Audio in the Bestselling Segment". In: *Journal of Cultural Analytics* 11, 1–30. 10.22148/001c.29802.

Berglund, Karl, Mats Dahllöf, and Jerry Määttä (2019). "Apples and Oranges? Large-Scale Thematic Comparisons of Contemporary Swedish Popular and Literary Fiction". In: *Samlaren* 140, 228–280. https://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-406939 (visited on 06/08/2026).

Björck, Staffan (1953). *Romanens formvärld: Studier i prosaberättarens teknik*. Natur & Kultur.

Cawelti, John G. (1976). *Adventure, Mystery, and Romance: Formula Stories as Art and Popular Culture*. University of Chicago Press.

Chatman, Seymour (1978). *Story and Discourse: Narrative Structure in Fiction and Film*. Cornell University Press.

Cleave, Chris (2008). *The Other Hand*. Sceptre.

Colgan, Jenny (2018). *Christmas on the Island*. William Morrow & Company.

Dahllöf, Mats (2022). "Quotation and Narration in Contemporary Popular Fiction in Swedish – Stylometric Explorations". In: *Proceedings of the 6th Digital Humanities in the Nordic and Baltic Countries Conference (DHNB 2022)*. Uppsala University, 203–211. https://ceur-ws.org/Vol-3232/paper17.pdf (visited on 06/08/2026).

Editeur (2025). *Thema current version 1.6*. https://www.editeur.org/151/thema (visited on 06/08/2026).

Fludernik, Monika (1993). *The Fictions of Language and the Languages of Fiction*. Routledge.

Folkhammar, Kristoffer (Aug. 2019). "Radioaktiv spänning i Sovjet". In: *Aftonbladet*.

Gelder, Ken (2004). *Popular Fiction: The Logics and Practices of a Literary Field*. Routledge.

Genette, Gérard (1980). *Narrative Discourse: An Essay in Method*. Trans. by Jane E. Lewin. Cornell University Press.

Håkansson, David and Carin Östman (2018). "'afbröt skolläraren ifrigt': En diakron studie av anföringssatsen i svensk skönlitteratur". In: *Samlaren* 140, 261–280. https://www.diva-portal.org/smash/get/diva2:1415008/FULLTEXT01.pdf (visited on 06/08/2026).

Herman, David (2006). "Dialogue in a Discourse Context: Scenes of Talk in Fictional Narrative". In: *Narrative Inquiry* 16 (1), 79–88. 10.1075/ni.16.1.11her.

Högström, Jesper (2013). "För smart för sitt eget bästa". In: *Dagens Nyheter*. https://www.dn.se/arkiv/kultur/for-smart-for-sitt-eget-basta/ (visited on 06/08/2026).

Jahn, Manfred (2021). *Narratology 2.3: A Guide to the Theory of Narrative*. University of Cologne. https://www.uni-koeln.de/~ame02/pppn.pdf (visited on 06/08/2026).

Jannidis, Fotis, Leonard Konle, Albin Zehe, Andreas Hotho, and Markus Krug (2018). "Analysing Direct Speech in German Novels". In: *Kritik der Digitalen Vernunft: Jahrestagung des Verbands Digital Humanities im deutschsprachigen Raum (DHd2018)*. 10.5281/zenodo.4622454.

Kacandes, Irene (2001). *Talk Fiction: Literature and the Talk Explosion*. University of Nebraska Press.

Kärnborg, Ulrika (Oct. 2014). "Spänning med guldkant av Carl-Michael Edenborg". In: *Aftonbladet*.

King, Stephen (2020). *On Writing: A Memoir of the Craft*. 20th anniv. ed. Hodder & Stoughton.

Kurfalı, Murathan and Mats Wirén (2020). "Zero-Shot Cross-Lingual Identification of Direct Speech Using Distant Supervision". In: *Proceedings of the The 4th Joint SIGHUM Workshop on Computational Linguistics for Cultural Heritage, Social Sciences, Humanities and Literature*. International Committee on Computational Linguistics, 105–111. https://aclanthology.org/2020.latechclfl-1.12.pdf (visited on 06/08/2026).

Lindemann, Charlotte (2019). "Dialogue and the Limits of Narrative Discourse: Gérard Genette, Gertrude Stein". In: *Word and Text: A Journal of Literary Studies and Linguistics* IX, 107–124. https://jlsl.upg-ploiesti.ro/documente/documente/Arhiva/Word_and_text_2019/05Lindemann.pdf (visited on 06/08/2026).

Malmberg, Carl-Johan (Nov. 2013). "Donna Tartt gestaltar konstens vidunderliga kraft". In: *Svenska Dagbladet*.

Mani, Inderjeet (2013). "Computational Narratology". In: *The Living Handbook of Narratology*. Ed. by Peter Hühn, John Pier, Wolf Schmid, and Jörg Schönert. Hamburg University Press. https://www-archiv.fdm.uni-hamburg.de/lhn/node/43.html (visited on 06/08/2026).

McIntyre, Dan (2010). "Dialogue and Characterisation in Quentin Tarantino Reservoir Dogs: A Corpus Stylistic Analysis". In: *Language and Style*. Ed. by Dan McIntyre and Beatrix Busse. Palgrave, 162–182.

Meister, Jan Christoph (2003). *Computing Action: A Narratological Approach*. de Gruyter.

Muzny, Grace, Mark Algee-Hewitt, and Dan Jurafsky (2017). "Dialogism in the Novel: A Computational Model of the Dialogic Nature of Narration and Quotations". In: *Digital Scholarship in the Humanities* 32 (2), 31–52. 10.1093/llc/fqx031.

Natt och Dag, Niklas (2017). *1793*. Forum.

Nivre, Joakim and Aaron Smith (2008). *UD Swedish Talbanken*. https://universaldependencies.org/treebanks/sv_talbanken/index.html (visited on 06/08/2026).

Peng, Qi, Yuhao Zhang, Yuhui Zhang, Jason Bolton, and Christopher D. Manning (2020). "Stanza: A Python Natural Language Processing Toolkit for Many Languages". In: *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. Association for Computational Linguistics, 101–108. https://doi.org/10.18653/v1/2020.acl-demos.14.

Phelan, James (2005). *Living to Tell About It: A Rhetoric and Ethics of Character Narration*. Cornell University Press.

Phelan, James (2011). "Rhetoric, Ethics, and Narrative Communication: Or, from Story and Discourse to Authors, Resources, and Audiences". In: *Soundings: An Interdisciplinary Journal* 94 (1–2), 55–75. https://www.jstor.org/stable/41200942 (visited on 06/08/2026).

Piper, Andrew (2018). *Enumerations: Data and Literary Study*. University of Chicago Press.

Piper, Andrew, Richard Jean So, and David Bamman (2021). "Narrative Theory for Computational Narrative Understanding". In: *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, 298–311. 10.18653/v1/2021.emnlp-main.26.

Rudberg, Denise (2019). *Under nio nätter* (*During Nine Nights*). Forum.

RWA (Romance Writers of America) (2025). *About the Romance Genre*. https://www.rwa.org/the-romance-genre (visited on 06/08/2026).

Sarenbrant, Sofie (2017). *Bakom din rygg* (*Behind Your Back*). Bookmark.

Squires, Claire (2017). *Marketing Literature: The Making of Contemporary Writing in Britain*. Palgrave Macmillan.

Sternberg, Meir (1982). "Proteus in Quotation-Land: Mimesis and the Forms of Reported Discourse". In: *Poetics Today* 3 (2), 107–156. 10.2307/1772069.

Thomas, Bronwen (2012). *Fictional Dialogue: Speech and Conversation in the Modern and Postmodern Novel*. University of Nebraska Press.

Todorov, Tzvetan (1977). *The Poetics of Prose*. Trans. by Richard Howard. Cornell University Press.

Vishnubhotla, Krishnapriya, Adam Hammond, and Graeme Hirst (2019). "Are Fictional Voices Distinguishable? Classifying Character Voices in Modern Drama". In: *Proceedings of the 3rd Joint SIGHUM Workshop on Computational Linguistics for Cultural Heritage, Social Sciences, Humanities and Literature*. Association for Computational Linguistics, 29–34. 10.18653/v1/W19-2504.

Wilkins, Kim, Beth Driscoll, and Lisa Fletcher (2022). *Genre Worlds: Popular Fiction and Twenty-First Century Book Culture*. University of Massachusetts Press.