

Haapanen, L., L. Kääntä & L. Lehti (toim.) 2018. Diskurssintutkimuksen menetelmistä. On the methods in discourse studies. *AFinLA-e. Soveltavan kielitieteen tutkimuksia 2018 / n:o 11*. 45–70.

Lauri Haapanen
University of Jyväskylä

Overcoming methodological challenges in text production research: a holistic approach through data triangulation

Text production research tends to analyse corpora of text products as its data. However, for the aim of investigating especially text production *processes*, such an approach falls short; a written text does its best to hide any traces of its genesis. This article argues for a holistic approach in text production research by presenting five methodological guidelines for future research: 1) what/how-research questions need to be followed by why-questions, and such research frameworks require 2) several methods to be applied in order to 3) encompass both product and process perspectives, 4) reveal material, mental and social activities, and 5) move from micro-level activities towards macro-level contexts. The holistic approach is empirically illustrated by drawing on a study on journalistic quoting (Haapanen 2017a).

Keywords: text production research, process vs. product, data, methodology, quoting

Asiasanat: kirjoittamisen tutkimus, prosessi vs. produkti, aineisto, metodologia, siteeraus



VERTAISARVIOITU
KOLLEGIALT GRANSKAD
PEER-REVIEWED
www.tsv.fi/tunnus

1 Introduction

A written text does its best to hide any traces of its genesis. Therefore, one cannot infer from the finished text, which is a *text product*, how it has come into being. For example, we do not know easily¹ whether a writer has typed the text linearly from beginning to end or revised the emerging text body by jumping back and forth.

This fundamental characteristic of written texts, together with the fact that writing has mainly been investigated from a product-bound perspective, thus neglecting procedural insights (for an overview, see Gresillón & Perrin 2014), have led to a limited understanding of real-life writing. The limitation results in approaches to text production processes that are sometimes speculative or even defective. For example, the much-cited insight that in the field of journalism “the higher the status of the speaker, the more direct the [quoted] presentation” (Davis 1985: 47) lives on (Bell 1991; Satoh 2001; Dai & Xu 2014). However, when scrutinised using data from different stages of production and adopting innovative methods which reconstruct journalistic decision-making, this interplay appears to be much more complex (Haapanen 2017b; Matsushita 2016).

The methodological dilemma in real-life text production research (that is, how to do empirical and analytical justice to the situated nature of writing in natural settings) largely derives from the challenges involved in data collection (Gresillon & Perrin 2014: 62). By investigating text production as a socially relevant activity of language use, in this article I argue for a holistic approach in text production research.² I suggest four data-related advancements that broaden the focus of analyses from products of writing to, firstly, material activities; secondly, mental activities; thirdly, social activities; and finally, contextual factors and traces that enable and constrain all the activities of writing. Data triangulation is a significant factor in this approach (as further explained in section 2), as I illustrate on the basis of my study on journalistic quoting (sections 3 through 7). The paper goes on to present some methodological guidelines for future research when reaching beyond the product-bound perspective (section 8). It concludes by making a point for innovative writing research in today’s technologised and mediatised society (section 9).

1 Of course, processes leave their traces in the final products, such as coherence gaps (e.g. Perrin 1999). With a certain expertise, text production researchers are able to draw conclusions from the shape of a final text product about the course of the writing process.

2 I am grateful to anonymous reviewers and to Daniel Perrin for our several inspiring discussions about the topic and his indispensable comments on the manuscript.

2 Four methodological steps towards a holistic view on writing

Capturing writing processes has long been methodologically challenging. Grésillon and Perrin (2014) have described this key problem in the history of applied linguistic-informed writing research with the help of *a double black box metaphor*. In this metaphor, the first of the boxes refers to the fact that real-life writing activities are performed behind the scenes, out of researchers' reach. Then, if researchers succeed in shedding light on these backstage processes, they are confronted with the second, inner black box, which is the fact that the material activity of writing hardly offers any certain information with regard to mental activities, so it is methodologically difficult to "open a window into the mind of the writer" (Perrin 2003: 915).³

In the next four subsections, I define four methodological steps towards a holistic view on writing. By doing so, I address the methodological challenges as represented in the double black box metaphor. Starting from the material activities of writing (section 2.1), I proceed to the mental activities (2.2), the social activities (2.3), and the contextual factors (2.4). Whereas the material activities are directly observable, the mental activities have to be inferred, for example based on writers' verbalisations. Social activities and contexts are partly observable, partly they must be deduced from both writers' activities and their verbalisations.

2.1 Material activities of writing

In order to overcome their products-only perspective and move, first, towards the material activities of writing, researchers need to access, for example, draft versions from different stages in writing (e.g. Grésillon 2014) or the source text on which the finalised product is based (e.g. the study on journalistic quoting, see section 3). This is what has been termed *version analysis* in media linguistics (e.g. Perrin 2013: 62). Due to the digitalisation of writing environments, more detailed data can be collected without interfering in the writing processes. Computer logging software running behind writers' text editors records all of one's screen activity and allows researchers to analyse all the keystrokes, mouse movements, pauses while writing and/or switches between applications (e.g. Baaijen et al. 2012; Beauvais et al. 2014;

3 According to widespread assumptions, activities, e.g. in writing, are either material or mental. However, it can be argued that mental activities such as reading and understanding change our neural states or even connections in our brains. By doing so, the mental activity results in a material change within our mental processor. Therefore, there are good reasons to assume that there is no such thing as immaterial, purely mental activities. Conversely, it can be argued that writing a word on a computer does not only cause material changes on screens and in brains; in an artificial intelligence framework, computers learn from this word, which can be considered an artificial mental activity.

Latif 2008; Morgan et al. 2013; Spelman Miller 2006; Perrin 2003). These data can also be further processed, for example, into graphs that visualise overall writing patterns (Fürier 2018; Leijten & Van Waes 2013).

2.2 Mental activities of writing

Tracking material activities falls short of analytical depth as soon as the research aims to do more than just *describe* the process of writing. If the process is also to be *explained*, researchers are confronted with the fact that the emergence of ink graphemes on paper and pixels on screens does not allow the analysis to explain why writers did what they did. A key in moving from material activities to mental activities of writing is to switch from *etic*, the research-analyst's perspective, to *emic*, the writer's insider perspectives.⁴ In practice, researchers need to ask writers to verbalise their thinking and decision-making.

Since writing is a highly verbal activity, methods for verbalisation that take place during the process of writing, such as a think-aloud protocol, most likely interrupt and distort the writing flow. They require verbal and metalinguistic activity that results in their interfering with and cannibalising the process under investigation (e.g. Jakobsen 2003; Smagorinsky 1989; Stratman & Hamp-Lyons 1994). Therefore, the verbalisation should preferably be performed afterwards. For relatively short writing assignments, the state-of-the-art example of such a method is cue-based retrospective verbalisation. Here, immediately after completion, writers view and comment on their video-recorded writing session. They thereby provide researchers with information about their awareness of what they are doing and why in the recording of the writing process. (E.g. Hansen 2006; Perrin 2003, 2013; other methods for retrospective verbalisation, see, e.g., Haapanen forthcoming; Leijten & Van Waes 2013; Manninen 2017; section 5.) However, when analysing and drawing conclusions from such data corpora, possible shortcomings must be taken into consideration. As Grésillon and Perrin (2014: 62) point out, "the access to the mind remains indirect, based on subjects' own explanations about what they have in mind and are able and willing to share with the researchers" (see also section 5).

2.3 Social activities of writing

By observing material activities of writing, we can see that some texts are produced in similar ways resulting in similar products, whereas other texts produced in seemingly the same kind of settings are different in terms of both products and processes (e.g.

4 As automated systems will increasingly be able to produce natural language, the concept of a *writer* also needs rethinking (see, e.g., Leppänen et al. 2017). However, in this article the automated generation of text and its reflections on writing research are not discussed.

Chen 2009; Fang 1991). Similarly, we can watch a group of conference attendees say hello to each other and observe that the men are shaking hands with other men but not with the women present. We can then ask people why they write specific text or greet peers the way they do. There will be answers – but most of these answers lack information about social patterns the individuals are not aware of. This particularly applies both to social taboos and to matters of course. Investigating as rigorously as possible the social components of activities such as networking or, in this case, writing, requires a blend of direct and indirect methods.

Whereas in the very beginning of text production research writing had been considered a product, and sometimes still is, a first paradigm shift resulted in a focus on processes. However, processes were conceived as changes on the mental and material levels only. A major reason for this epistemological limitation can be seen in the fact that empirical research into writing processes was launched and long dominated by psychologists, who investigated writing as a set of decontextualised mental and material activities in laboratory experiments. It was only through ethnographical text production research, in a second paradigm shift in the field, that the focus then moved towards writing as a social practice. (For a brief overview, see, e.g., Perrin 2013: 53–54; for a forerunner in the field, see Hodge 1979.)

By taking “a social perspective on writing in the workplace”, text production research elaborated on the concept of collaborative text production in “communities of discourse” and “communities of practice” (Pogner 2012). This means that, for example, news is not produced by individual authors but in processes based on the division of labour. Journalists cooperate with co-contributors such as sources, peers, photographers, video-editors, copy-editors, and audiences providing feedback through e-mails, phone and social media. This co-construction in the co-production of news can be investigated through observation of social and material activity (such as e-mail exchange) and, again, through verbal data of people talking about their production processes and mentioning social interaction.

2.4 Contextual factors and traces of writing

Applied linguistics explicitly focusses on theoretically and empirically investigating “real-world problems in which language is a central issue” (Brumfit 1995: 27). From a practical perspective, this means identifying and explaining practical problems and contributing to their solution (AILA 2018). Of course, “[it] is hard to think of any ‘real-world’ problems – from global warming to refugees to generic counselling to outsourced call centres to AIDS/HIV to military intelligence – that do not have a crucial component of language use” (Myers 2005: 527). Moreover, in today’s mediated communication, it is hard to think of such problems that do not have a crucial component of writing.

As always in the interplay of micro and macro structures and processes, what people write is influenced by and influences the overarching real-life environment, for example its social structures, and vice versa (for an overview of integrative social theories and their relevance for text production research, see Perrin 2013: 215–222). This interplay is infinitely dynamic and complex (e.g. Bernstein 2015; Cilliers 1998; Perrin 2013: 235–245), which is why epistemological reductions must be conducted in order to identify and conceptualise the significant and important contextual factors and traces of writing. The next four paragraphs outline complementary approaches to reducing such complexity according to specific research questions.

Jakobs (2005) has created a context model of five nested ranges of interacting contextual factors, which both enable and constrain the process of writing. According to this model, writing is affected by

- 1) the *writer's* individual properties and personal preferences, etc;
- 2) his or her *workplace's* hierarchical aspects, budgets, and material conditions, etc.;
- 3) the *organization's* values, quality expectations, and communication policies, etc.;
- 4) the *domain's* standards, routines, and professional education, etc.; and
- 5) the *culture's* ethical norms and legal rules, etc.

An obvious methodological challenge is how to focus research on the factors that are contextually (the most) revealing in any given circumstance. To give but a few illustrations of innovative approaches concerning that aim, Wortham and Reyes (2015), for example, draw on linguistic anthropology and present a methodology that identifies and interprets deictics, reported speech and other indexical signs from parallel speech events, thus allowing researchers to understand not only individual events but also the discursive patterns that emerge between and from them.

Vandendaele et al. (2015), in turn, focused specifically on the craft of sub-editors in newsrooms to dispel the belittling belief that sub-editors merely “trim the fat” from overlong manuscripts. The researchers, first, exploited data from ethnographic observations of practitioners and did version analyses of unedited and edited manuscripts to show that the most powerful sub-editing work was addressed to features such as headlines, photo captions, pull-out quotes and bulleted fact-boxes. They then triangulated these findings with the results from readers’ eye-tracking tests to reveal that these features specifically were the “entry points” where readers stop skimming the text and start to read it. Finally, they interviewed the sub-editors to confirm that the focus on these important entry points was intentional on their side. (For triangulation of multiple methodologies, also see, e.g., Beaufort 2008; Grésillon & Perrin 2014; Van Hout 2015; Van Hout & Macgilchrist 2010; Velthuis 2016.)

Lillis (2008), in order to build a holistic understanding of academic writing, broadened the temporal dimension of the data triangulation. In practice, the research team engaged in the participants' academic writing world for years and collected multi-layered data in the form of article manuscripts, various types of interviews and correspondence, observation notes and a large number of photographs and institutional documents. Thus, the researchers aimed at *thick participation* – “a form of socialisation in order to achieve a threshold for interpretive understanding” (Sarangi 2006: 73) – that helped them “foreground what is significant to writers from their specific sociohistorical perspectives” (Lillis 2008: 373). (For a longitudinal research orientation, see also, e.g., Compton-Lilly 2014; Lammers & Marsh 2018; Saldaña 2003.) Taken together, the approaches presented in this subsection illustrate the close relation between research question, context and focus of epistemological reduction.

In the sections that follow (3–7), four advancements presented in this and the previous subsections (2.1–2.4) are operationalised in pursuing the process of quoting in written journalism. In practice, I walk my readers through a research design that illustrates a theoretically sound triangulation of methods for collecting data about material (section 4), mental (5) and social (6) activities as well as relevant contextual factors and traces of writing (7) to answer research questions triggered by a real-life problem (3). While the results are excluded from this presentation (for them, see Haapanen 2017a), the preparation for, and conduct of, the data collection and its analysis are explained in detail in order for readers to be able to evaluate the reliability and validity of the methods used.

3 The study: from a real-life problem to research questions

Research in applied sciences is typically triggered by a real-life problem. In my example, which draws on my research project on journalistic quoting (Haapanen 2016a, 2016b, 2016c, 2017a, 2017b, 2018), without explicit cross-references, the trigger was an evident contradiction that appeared between institutional meta-discourse about quoting on the one hand, and actual quoting practices described in existing research on the other. The former, presented in guidebooks, editorial policies and self-regulatory guidelines, presumes that the quoted text is a fairly exact reproduction of what a source used as a text agent has said – if not word-for-word, then at least in a meaning-for-meaning way. The latter, in turn, hints that the relationship between the original and the quoted discourse is by no means so simple to conceptualise (Haapanen 2017a: sections 1.2.1 vs. 1.2.2).

As a quote claims to render what someone has said, I conceptualised the process of making quotes as *recontextualisation*. In doing so, I followed Per Linell's (1998: 154) definition of recontextualisation as the "dynamic transfer-and-transformation" of some part or aspect from one discourse to another. The transferred elements may be, for instance, specific lexical items, arguments, narratives, values, conceptions, and/or ideologies. An interesting issue for research is the fact that through recontextualisation, the quoted discourse is subject to various changes (ibid. 155). My goal was to explore the relationship between the original spoken and the final written discourse by asking: **How are journalistic interviews recontextualised into written quotations in journalistic articles?**

Linell (1998: 154–155) suggests that the process of recontextualisation consists of three consecutive sub-processes: the discourse that is selected to be recontextualised is first extracted, then re-positioned, and finally modified to fit the context. I exploited and elaborated this three-part structure in the design of my research framework, and named the sub-processes *decontextualisation*, *contextualisation* and *textualisation*. Within the recontextualisation of utterances from interview to quotes, the quoted piece of discourse undergoes various changes. Regardless of heterogeneity, there are reasons to assume that these changes reproduce and renew some common patterns, which are routines on the process level and genre elements on the product level. Otherwise, there might be severe delays and interruptions in the performance of a journalist's daily tasks.

To formulate the second research question, I conceptualised the institutionalised transformation from interviews to articles as an *intertextual chain*, a term coined and defined by Norman Fairclough (1992). During this "chaining", a particular type of text is transformed into another type of text "in regular and predictable ways" (ibid. 130), so my goal was to trace the regular and predictable ways that guide the process of formulating quotes. I did so by asking: **How can we explain those quoting strategies⁵ that link the original interview discourse to the final quote discourse?**

In the light of these research questions, I decided on some methodological principles as the basis for the research design.

- I prefer empirical **field studies** over laboratory research, because the field of activity is largely unexplored but most likely so complex that it cannot be made the focus of an experiment with a restricted number of strictly

5 Writing strategy refers to "the reinforced, conscious, and therefore articulable idea of how decisions are to be made during the act of writing so that the writing process or text product has a great probability of fulfilling the intended function. Strategies are recursive: they can contain (sub)strategies. Individuals and collectives have repertoires of writing strategies: the sets of strategies available when writing." (Perrin 2013: 55.)

controlled parameters (cf. e.g. Beers et al. 2017; Mangen et al. 2015; Olive & Barbier 2017).

- I decide for generalisation based on **theoretical sampling** instead of **statistics**, since answering the research question requires in-depth analyses of case studies. According to the principles of grounded theory (e.g. Glaser & Strauss 1967; Tavory & Timmermans 2009), the cases are selected and analysed in cycles until theoretical saturation.
- I exploit both **observations** and **interviews**, because the former captures informant-journalists' actual activities but leaves it to the researchers to interpret their intentions, while the latter allows researchers to access informant-journalists' mental reflections but is constrained by their self-awareness (cf. e.g. Perrin 2013).

4 Material activities: linguistic modifications in quoting

In order to access journalists' material activities of quoting, I needed data to consist of (at least) the relevant extracts from original and final discourse, namely oral interviews and published articles with one or more direct quotes.

To begin with, there were no research results indicating that a specific article type (e.g. news, profiles, fact-focused articles), publication (some particular title) or media genre (e.g. national/regional newspapers, women's magazines, bulletins) would be the pivotal factor so that, for example, in the production of sports magazines, spoken discourse would be converted into written quotes in some systematic way. Therefore, I considered **the journalistic interview** to be an appropriate starting point for my data collection as it is a central and conventionalised premise for information gathering in journalistic work (Ekström 2006: 23).

An initial (and then, in terms of grounded theory and theoretical sampling, ongoing) task was to select and contact prospective informant-journalists. The key feature of the selection throughout the process was that they regarded themselves as professional journalists and were employees or freelancers for established media that published their content in Finnish. Additionally, I aimed for a diverse set of journalists in terms of gender, age and the type of publication for which they worked. Only six of the journalists I contacted refused to participate in my research. This relatively fortunate result (cf. Bell & Garrett 1998: 19) might have been because I emphasised to the prospective informant-journalists that I was describing, not reviewing or evaluating, their work performance. I did not disclose at this point the

exact objective of my research in order not to intrude in their work. Of course, this objective needed to be explained right after the data collection.⁶

After selecting the informant-journalists, I asked them to record their **journalistic interviews (= data set 1)** for my research. In most of the cases, I covered one text production process per informant-journalist, because awareness of the data collection methods, especially of stimulated recall (see section 5), might have influenced the workflow of the following case. In only a few cases did I cover two processes in order to examine the quoting strategies of an individual journalist on different assignments. In these cases, both stimulated recall sessions were recorded together, upon completion of the second of the two text production processes (for similar procedures when working with large data corpora, see Perrin forthcoming).

To collect this data, video recording would very likely have provided valuable instances of nonverbal communication during the interview. However, videotaping might have had an undesirable influence on the interview by distracting both the interviewer and the respondent. An audio recording, in contrast, worked better, as it is a commonly used procedure in the field of journalism.

I also recorded some press conferences for my data, because they are, like one-on-one interviews, a common way of collecting data for articles. One advantage of this decision was that it makes my research more comparable to existing studies on quoting in written journalism, since they only contain examples of press conferences (Johnson Barella 2005; Lehrer 1989).

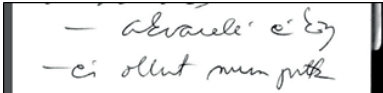
From those of the journalists who took notes by hand during the interview (and made audio-recordings only for my research purposes), I also asked permission to copy their notebooks. I refer to these texts as **transitional text documents (= data set 2)**, because they were made to facilitate the conversion of the original interview discourse into the target article text.

After collecting the interview recordings and transitional text documents, I collected **the published articles (= data set 3)** that were written based on these interviews. The articles contain 165 quotes altogether, ranging from 3 to 26 quotes per article. The articles were published in newspapers, magazines, business-to-consumer magazines/bulletins and online-publications, and they were collected in their original layout, so the possible influence of layout and certain visual elements could also be taken into account in the analysis.

Table 1 summarises the data collection for tracking linguistic modifications during the recontextualisation process. I also give a sample of the data using an example referred to as "Painting with oils". In this case, the journalist provided me with a digital recording of an interview with a local artist who paints sea motifs. She took notes for herself on a thin notepad – the recording was made and used only for

6 The journalists had the opportunity to pull out at this moment, despite their initial decision to participate – all of them stayed on board.

my research purposes – and this notepad is also accessible below, in addition to the interview and article texts.

Recordings, transitional texts and published journalistic articles		
Description	Content in total	Example "Painting with oils"
Recordings of oral journalistic interviews and their transcripts	20 recordings <ul style="list-style-type: none"> • 17 one-to-one interviews • 3 press conferences 	<p>Example in Finnish (original)</p> <p>JO = Journalist IN = Interviewee</p> <p>JO: maalaat öljyväreillä IN: joo JO: eikö öö tällasta merellisyyttä olisi helpompi toteuttaa vaikka akvarelleilla IN: ei JO: miksei IN: niinku mä sanoin just mä meen paperin läpi (.) ei käy ei käy (.) ja sit se on se et' tota (.) joo se ei vaan sovi mulle (.) mä oon joskus kokeillut kauan kauan sitten mut (.) se ei oo mun juttu (.)</p> <hr/> <p>Example in English (translation)</p> <p>JO: you paint with oil colours IN: yes JO: wouldn't it be easier to depict such maritime motifs with, for example, watercolours IN: no JO: why not IN: like I just said, I'll tear through the paper (.) no way no way (.) and then the thing is that (.) yeah, they just don't suit me (.) I tried them a long long time ago but (.) it's not my thing (.)</p>
Transitional text documents	5 sets of hand-written notes	 <p>– akvarelli ei käy – ei ollut mun juttu</p> <hr/> <p>– watercolours don't work – was not my thing</p>

Published articles	<p>21 published articles (165 quotes) by 17 journalists.</p> <ul style="list-style-type: none"> • 17 based on interviews • 4 based on press conferences (2 based on the same press conference) 	<p>Example in Finnish (original)</p> <p>Sofia maalaa öljyväreillä. Niillä hän loihitii esiin aallot, pärskeet ja meren vellovan voiman, mutta maalauksissa on myös akvarellimaista herkkyyttä. Akvarelleja hän ei voi kuvitella maalauksiansa.</p> <p>“Olen liian temperamenttinen. Paperi meni puhki yhdellä siveltimenvedolla”, hän nauraa.</p>
		<p>Example in English (translation)</p> <p>Sofia paints with oil colours. With them she conjures forth the waves, the splashes and churning power of the sea, but the paintings also have a sort of watercolour-like delicacy. She cannot imagine painting with watercolours.</p> <p>“I’m too temperamental. The paper would be torn with a single brushstroke,” she laughs.</p>

TABLE 1. Data collection for the tracking of linguistic modifications. (Data are originally in Finnish, translated into English by the author.)

The analysis of these data sets was not particularly focused on any of the sets per se, but on the differences between them. Therefore the overall analytical method exploited to analyse such data is called version analysis (applied in media linguistics, see Perrin 2013: 62). As a method, version analysis is suitable for reconstructing the changes that linguistic features undergo from one textual version to another. In my research located in the Finnish mediascape, the basis for comparing versions is text analysis that exploits the premises and research results of (Fennistic) interactional linguistics, discourse studies, as well as lexicology and the study of grammar.

5 Mental activities: tracing journalists’ quoting strategies

A conventional option for capturing the journalist’s mental activities of quoting would have been to conduct an interview-based inquiry (Grésillon & Perrin 2014).

However, such an approach could cast doubt on the credibility of the results. One such pitfall materialised in my earlier research (Haapanen 2011), when I conducted a series of semi-structured interviews with experienced journalists and journalism educators inquiring about their perceptions of quoting. The answers I received mostly repeated the ideal procedures presented in journalism guidebooks (see also Mitchell & Rosenstiel 2000). It is possible that journalists were not fully aware of their actual quoting strategies and practices. It is possible, too, that journalists consciously dressed up some of their practices that breached the general guidelines. Furthermore, given the fact that quoting is such an everyday process, presenting questions about it is as if the researcher were “outsourcing” the task of generalisation to the informant, who then does it without following any consistent methodological procedure. Based on these observations, I argue that semi-structured interviews have a chance to produce reliable knowledge only when applied to one-off events as soon as possible after their occurrence.

To overcome the potential problems, an interview-based inquiry needs to be linked concretely to the parts of the intertextual chain, i.e. the interview and (the writing of) the article, to provide a better insight into the actual work practices of journalists (see also Androutsopoulos 2008: 8, and Perrin 2003: 915–917). I therefore adopted the **stimulated recall (SR)** method (DiPardo 1994; Lindgren & Sullivan 2003; for an overview, see Lyle 2003). Traditionally, SR begins with videotaping a selected person at work, after which the person is asked to view and comment on the video. The method is designed to increase people’s awareness of their performance and thus help them reconstruct the trains of thought they had while working. Due to the stimulus or cue (traditionally the videotape), the method also prompts informants to “discuss processes and interactions that they otherwise might have neglected” (Smagorinsky 1994: xv; see also Dempsey 2010: 350–351). In my application of SR, I used the transcript of the recording of the original journalistic interview (data set 1) and the published article (data set 3) as stimuli for the verbal protocol (**data set 4**) that reconstructed the informant-journalist’s quoting process.

Each SR session consisted of four elements. The session began with some questions about the informant-journalist’s **biographical and background information**. At this stage, I also explored the informant’s general perceptions and rules of thumb about quoting. As mentioned above, my aim was to include a diverse set of informant-journalists. The accumulation of biographical information proved that besides gender, age and the type of publication worked for, the informants’ work experience and professional education (or lack of it) also varied considerably and reflected well the general diversity of journalists’ backgrounds.

The aim of the **first cycle of SR** was to determine the strategies used with regard to the first and second sub-processes of quoting, namely decontextualisation and contextualisation. In essence, I asked the informant-journalist to look at the published article and explain, first, why s/he had selected this particular text segment

or content to be quoted, and second, why the quote was positioned in that particular place in the article. During the **second cycle of SR**, the informant-journalist and I closely read the published article and the transcript of the journalistic interview. The main objective of this cycle was to determine the strategies for the textualisation of the quoted discourse. Alongside this formal procedure, my SR sessions aimed to reveal **other quoting-related issues**, such as the factors that influence quoting practices as well as metadiscourse that evaluates his/her quoting. In this way, I was able to reconstruct the individually- and socially-anchored language awareness in a discourse community (with the latter I refer to the review of the written institutional guidelines).

Altogether, I conducted SR sessions with 11 journalists. At this point, the answers appeared to be “saturated” in that significantly new aspects no longer arose during the last sessions. The following table (Table 2) summarises the content of a SR session, which will also be illustrated by excerpts drawn from the example “Painting with oils”.

After collection, the SR data were analysed by exploiting a **two-cycle qualitative content analysis procedure** proposed by Johnny Saldaña (2009). In the first cycle of the analysis, I identified similar quoting processes which emerged from the retrospective protocol data (SR Cycle 1 and 2). I then labelled the segments with shared features with a common code, for example, in terms of topic, purpose, goal or practice. I will illustrate this analysis with the example “painting with oils” (see also Table 3): the segment that is relevant to the decontextualisation phase of the quoting process (*“It was difficult to find good quotes, because...”*, see Table 2) was assigned the code CHARACTERISING THE SPEAKER AND DESCRIBING THE INTERVIEWEE’S WAY OF SEEING THE ISSUE. The segment pertaining to the contextualisation phase (*“I don’t plan the structure of an article in advance but...”*), was, in turn, assigned the code CONSTRUCTING THE NARRATION OF THE ARTICLE. The segments of the SR that concerned the textualisation phase of the quoting process were then attributed the codes CLARIFYING THE QUOTED DISCOURSE (*“She spoke so quickly and so much that I had to shorten and clarify this quotation.”*) and MODIFYING INTO STANDARD LANGUAGE (*“I standardised her colloquial speech.”*).

Four elements of a SR session		
Phases of the SR	Goal of the phase	Example "Painting with oils"
Biographical and background information	To learn the informants' ages and educational backgrounds, work histories, as well as the guidance they have received on quoting, if any, etc.	Middle-aged woman, unfinished university studies (Journalism as a minor), a couple of 2–3-day update training courses.
		More than 30 years of work experience in regional and local newspapers. Mainly as a news reporter. Cannot recall any instructions on quoting from either guidebooks or the editorial offices where she has worked. Always takes notes by hand. With controversial topics (e.g. politics), she sometimes also tape-records the interview to protect herself from possible repercussions.
Two cycles of close reading	To determine the practices of the decontextualisation and contextualisation phases	"It was difficult to find good quotes, because this interviewee didn't analyse her own work practices as profoundly as some other artists." However, "she spoke strongly with her hands, and it was interesting. Here she showed with her gestures that the paper would tear", so the journalist decided to decontextualise this aspect of the interviewee's turn. "I don't plan the structure of an article in advance but instead I often pre-arrange a list of questions which then steer the line of the emerging article", as happened this time as well. It's worth mentioning that sometimes some good quotes – "pearls" as she called them – need to be positioned at the beginning of the article, and in this way quotations clearly affect the whole structure of the article.
	To determine the practices of the textualisation phase	"She spoke so quickly and so much that I had to shorten and clarify this quotation. (...) I standardised her colloquial speech."
Other quoting-related issues	Influencing factors on quoting, the role of quoting in the writing process, evaluations of informant-journalists' own quoting practices, etc.	"In my opinion, a quotation must be short and snappy. I don't know why... maybe it's more reader-friendly that way. The best bits as quotations and the rest somehow inserted into the text, modified in one way or another." <i>Do you discuss quoting practices in some particular assignment or generally?</i>
		"No. When we assemble a newspaper issue we are in such a great hurry that there's no time. I would love to discuss work practices, but our schedules are so tight that there's no time for that."

TABLE 2. The course of a stimulated recall (SR) session. (Data are originally in Finnish, translated into English by the author.)

In the second cycle of the analysis, I identified the similarities in the patterns of quoting processes and organised similar codes into categories. The goal of this second cycle was to reach a compact and credible number of conceptual categories in relation to the range of the retrospective verbal protocols collected in the SR sessions. I will again illustrate this second cycle of the analysis with the example “Painting with oils”: The first two aforementioned codes fell into the category *CONSTRUCTING THE PERSONA OF THE INTERVIEWEE*. The third code, dealing with contextualisation, was linked with another similar code, and they were labelled *CONSTRUCTING THE NARRATION*. The codes relating to the textualisation phase – *CLARIFYING THE QUOTED DISCOURSE* and *MODIFYING INTO STANDARD LANGUAGE* – were positioned within two categories, *CLARIFYING THE ORIGINAL MESSAGE* and *STANDARDISING THE LINGUISTIC FORM*, respectively.

In its entirety, I identified nine categories covering the strategies that create intertextual chains between the journalistic interview and quotes. These strategies were then connected axially to detect a *core category* (Saldaña 2009: 163–167). The core category is composed of all the products of the analysis and is condensed into a few words. It is illuminating for the analysis as it “explain[s] variation as well as the main point made by the data” (Strauss 1987, quoted in Strauss & Corbin 1998: 147). The core category abstracted from the SR data is *EXECUTING THE OBJECTIVE(S) OF THE ARTICLE OVER THE DEMAND FOR “DIRECTNESS”*.

An illustration of the analysis with “painting with oils”				
Phase of the quoting process	Excerpt from data set 4	Coding → (1st cycle of the analysis)	Categorising → (2nd cycle of the analysis)	Core category
Decontextualising	It was difficult to find good quotes, because...	CHARACTERISING THE SPEAKER DESCRIBING THE INTERVIEWEE’S WAY OF SEEING THE ISSUE	CONSTRUCTING THE PERSONA OF THE INTERVIEWEE	EXECUTING THE OBJECTIVE(S) OF THE ARTICLE OVER THE DEMAND FOR “DIRECTNESS”
Contextualising	I don’t plan the structure of an article in advance but...	CONSTRUCTING THE NARRATION OF THE ARTICLE	CONSTRUCTING THE NARRATION	
Textualising	She spoke so quickly and so much that I had to...	CLARIFYING THE QUOTED DISCOURSE	CLARIFYING THE ORIGINAL MESSAGE	
	I standardised her colloquial speech.	MODIFYING INTO STANDARD LANGUAGE	STANDARDISING THE LINGUISTIC FORM	

TABLE 3. An illustration of the course of the analysis.

6 Social activities: reconciling the expectations of the employer, interviewees and readers

Through the retrospective verbalisation produced in the stimulated recall (SR) session of the “Painting with oils” case, we can also reveal how the journalist customised her linguistic product to meet her organisation’s requirements, her interviewee’s expectations and the target audience. Conversely, we see how the journalist perceived the influence these social agents exerted on her work and how they were negotiated according to her memory. In addition, and still based on verbal data, we can reconstruct her social contacts related to the text production.

For example, the journalist told the researcher that “for the artist [= interviewee], it seemed to be very important to get publicity for her upcoming exhibition”. As she was working for a local newspaper funded by advertising, the journalist knew that her employer was generally inclined to offer such publicity to local agents, such as this artist.

However, meeting these social expectations was a tricky business because the interviewee “didn’t analyse her own work practices as profoundly as some other artists”. For this reason, the interview material lacked the sort of statements that are especially quotable, namely statements that reveal interviewees’ mindsets and the way they structure their thoughts (Haapanen 2018: 108). Nevertheless, all the interview-based articles in this newspaper – and, nowadays, in written journalism generally – should include quotes, and those quotes “must be short and snappy” in order to be “reader-friendly”, the journalist explained in the SR session.

In this assignment, the journalist did her best to fulfil all the social expectations of the employer, interviewee and the readers. However, sometimes the negotiation of these expectations also offers the journalist an opportunity to show her stance while remaining ostensibly neutral – as journalists are often thought and required to be. For example, the journalist looked back on a case in which she put someone whom she considered to be an arrogant interviewee in his place: a politician she disliked made important points in the interview which she, then, disseminated in her news item, but as general narrative, without attributing them to that politician.

Although recent research has emphasised the interactional nature of journalistic work, contrary to the traditional idea of journalists as lonely wolves (e.g. Merminod 2018; Perrin 2013), the journalist who wrote the case “painting with oils” actually longed for professional interaction, which she said had decreased due to austerity measures: work processes were streamlined with the help of publishing system software that determines the fixed parameters of articles, thus removing the need to collaborate with layout editors; sub-editors and superiors were too busy to discuss and give feedback on manuscripts; the number of in-house photographers had been reduced and journalists mainly took the photographs for their articles themselves; the employer did not encourage employees to participate an in-service

training anymore. The discontent came across from the verbal protocol, but it was also easily observable for the researcher in the SR session from the journalist's facial expressions and gestures.

7 Contextual factors and traces: investigating quoting in the light of media concepts

To relate the material, mental and social processes of writing with the contexts of production in which they take place and, thus, by which they are influenced, a more complex research frame was needed. I aimed to explain the modifications revealed by the version analysis of data sets 1, 2, and 3, for which the informant-journalists gave their reasons in the stimulated recall sessions (data set 4), through the notion of *media concepts*.

The notion of **media concepts** introduces and organises the relevant extra-linguistic contexts and contextual resources that affect the process of creating any specific media product. The notion is based on cultural-historical activity theory. This theory conceptualises organisations as activity systems which have historically and socially developed goals and purposes; in activity theoretical terms, every organisation has an object of activity which materialises in some outcome – either as services or, as in my research topic, as a product (Engeström 1987). The key point of activity theoretical thinking, in the light of my research, is that within the activity system, the work practices of a practitioner are not merely an individual or independent piece of craft, but are influenced by the external and internal contexts of the work process.

However, activity systems are always heterogeneous and multi-voiced, because different practitioners construct the object and the other components of the activity in different, partially overlapping and partially conflicting ways (Engeström 1987). For example, the journalist, editor-in-chief, graphic designer, and advertising sales representative work towards the same final outcomes, which include a good newspaper, but on the micro level they most probably have different ideas about what is "good". Such internal tensions in the activity system appear as disturbances, such as errors, problems, breakdowns, and ruptures in communication. Therefore, activity theory and especially developmental work research analyse disturbances as a window to deeper structural and historical contradictions in work practices within the activity system and a network of activity systems (Helle 2000).

Based on activity theory and developmental work research, Jaakko Virkkunen has developed *the concept of activity* to model and describe activity systems (e.g. Virkkunen 2007). The concept of activity consists of three components: 1) the purpose and values of an activity system, 2) the artefact or service produced, and 3) its production in daily practice. By identifying these components, the concept of

activity makes it possible to better coordinate individual actions as the concept is “embedded in the structures and daily practices of the activity” (Virkkunen 2006: 46).

The notion of a media concept was formulated by Finnish journalism scholars Merja Helle and Maija Töyry (Töyry 2005, Helle & Töyry 2009, Helle 2010). It is an adaptation of the concept of activity to media research, and it is structured into three mutually constitutive and closely intertwined components, which organise contextual resources that affect the production of any specific media product (see Figure 1; for more detail, see Haapanen 2017a: 37). Within my research design, I use the notion as empirically grounded modelling to relate a single quoting practice to the broad object of the activity system.

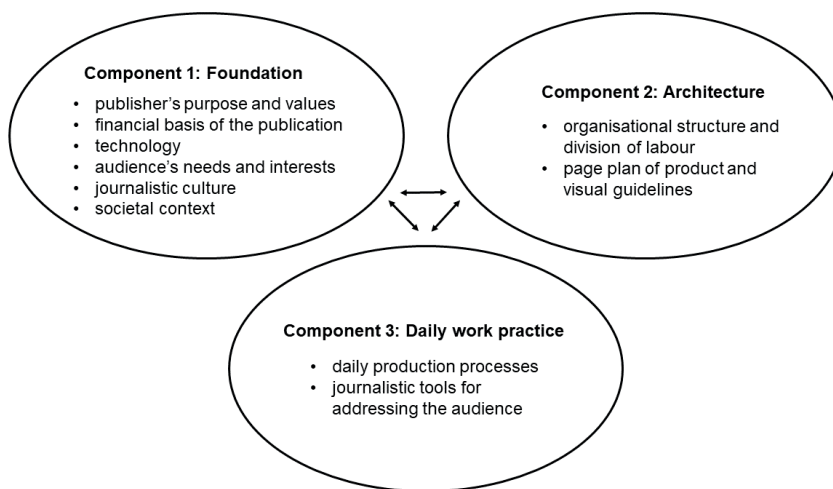


FIGURE 1. Components of a media concept (based on Helle & Töyry 2009: 502).

The stimulated recall sessions as well as the analysis of the institutional meta-discourse brought to light various points and factors related to quoting. In this phase, each of these aspects was placed under one of the three components of a media concept. Thus, with regard to the second research question (*How can we explain those quoting practices...?*), the notion of media concepts helped to explain single quoting-related aspects as interdependent constituents in a broader context of journalistic article production.

The analysis cast light on the multi-dimensional interplay of the various contextual resources that take place in quoting, and revealed that quoting is a process of constant internal negotiation between various points and aspects that might be difficult to reconcile (e.g. Zampa & Perrin 2016). Furthermore, these aspects may originate from different fundamental premises and be determined by different stakeholders. For example, inadequate resource allocation (the organisational

architecture, Component 2) may cause time pressures that, in turn, lead to the selection of disadvantageous working methods (daily production processes, Component 3), such as taking notes by hand because there is no time for tape-recording and transcribing, despite its inaccuracy. However, whereas the commonly mentioned goal of reproducing the original idea of the interviewee's utterances in quotes (journalistic culture, Component 1) may suffer from haste, the technical ease of modifying the quoted text (technology, Component 1) can help the writer cope with insufficient raw material. At the grass roots level, these disadvantages must be managed by the journalists themselves, although internal contradictions may originate from fundamental premises that individual journalists can neither affect nor change (for handling contradictory expectations through journalistic practices, see also Perrin 2013: 33–35, and Zampa & Perrin 2016: 17–19).

Certainly, taking into account the increasingly complex contexts in which the process of quoting is embedded, an all-encompassing perception of omnipresent and multi-layered contextual resources is impossible to achieve. However, the notion of media concepts offered an empirically grounded model to relate a single observable, material practice or cognitive strategy to the activity system and to other broader contexts.

8 Methodological guidelines for adding value to text production research

The research on journalistic quoting succeeded in describing and explaining the contradiction between institutional meta-discourse and actual quoting practices. It also revealed substantial and fundamental contradictions within the process of quoting.

To encourage a holistic approach to future writing research in the field of applied linguistics, in this section, I will outline five methodological guidelines that reflect the advancements that were presented in section 2. These guidelines are not offered as dichotomies, but as complementary groups in which each element supplements the others' potential. *WHAT/HOW*-research questions need to be followed by *WHY*-questions (1), and such research frameworks require several methods to be applied (2) in order to encompass both product and process perspectives (3), reveal material, mental and social activities (4), and move from micro-level activities towards meso- and macro-level contexts (5).

- 1) **WHAT/HOW-research questions** describe the phenomenon under study. However, if one aims not only to *describe* but also to *explain* the observable outcome – and, thus, increase the potential for creating added value also for practitioners by bringing out unconscious and automi-

sed, imperfect work activities – these questions need to be followed by **WHY-research questions**.

- 2) Data collection and analysis can be conducted through a **single and (relatively) simple method**, for example, by producing a data corpus by selecting finalised text products and then analysing this corpus with a method of (a kind of) text analysis. However, (adequately) reconstructing the material and mental complexity of real-life writing, and connecting situated activities within social structures and processes, probably require the application of **several methods** of data collection and analysis.
- 3) **Product perspective** allows researchers to make only informed guesses concerning how the text has come into being. A **process perspective**, in turn, also requires capturing the process of writing character by character or by comparing two or more consecutive text versions.
- 4) In order to explain observable **material activities**, sophisticated methods for prompting writers to verbalise their **mental** and **social activities of writing** must be adopted. Writers' emic perspective also provides research with information about what is or is not significant from the infinite range of contextual resources.
- 5) On the **micro level**, writing research investigates phenomena in their immediate co-texts and contexts. However, since real-life writing takes place in real-life circumstances instead of a vacuum or laboratory experiment, considerable explanatory power can only be achieved by situating these activities of writing (micro level) in layered contexts, such as concrete workplace environments (**meso level**) and their (not directly accessible, yet influential) institutional, societal and cultural contexts (**macro level**).

9 Conclusion

For a long time, writing research has neglected procedural insights and writers' emic perspectives. This shortcoming stems from the challenges in data collection as well as from the dominance of product-bound text analysis in linguistics and sometimes simplistic interview methods in the social sciences, and it has resulted in constricting the picture of real-life writing. Drawing on state-of-the-art research, this article sought to overcome these methodological challenges by laying out guidelines for a holistic approach to applied linguistic-informed writing research.

There is clearly a call for writing research. In today's technologised and mediatised society, writing is omnipresent in education and working life as well as increasingly in everyday life, especially through social media (Bazerman 2009; Jakobs & Perrin 2014). Furthermore, the conventional way of producing texts, *focused writing*, in which writers revise and polish their texts until close to perfection,

now works in conjunction with *writing-by-the-way*, which is instant and incidental by nature and transgresses boundaries between spontaneous speech and well-groomed text (Hicks & Perrin 2014).

To answer this call for writing research, modern technology is on our side. Digital writing environments enable researchers to track writing as never before, from various perspectives and levels and in a relatively non-invasive way. However, non-digital activities and the contexts in which the process of writing is embedded are still challenging to capture in their dynamics and complexity.

To conclude, in the field of applied linguistics-informed writing research, a lot of good work has already been accomplished but a lot still remains to be done. Innovative methods and methodologies must therefore be further developed, applied, and evaluated.

References

- AILA [The International Association of Applied Linguistics] 2018. *What is AILA – mission and activities*. <https://aila.info/about/> [accessed 24/8/2018].
- Androutsopoulos, J. 2008. Potentials and limitations of discourse-centred online ethnography. *Language@Internet*, 5, article 8. <http://www.languageatinternet.org/articles/2008/1610>.
- Baaijen, V. M., D. Galbraith & K. de Glopper 2012. Keystroke analysis: reflections on procedures and measures. *Written Communication*, 29 (3), 246–277. <https://doi.org/10.1177/0741088312451108>.
- Bazerman, C. (ed.) 2009. *Handbook of research on writing: history, society, school, individual, text*. New York, NY: Lawrence Erlbaum Associates.
- Beaufort, A. 2008. Writing in the professions. In C. Bazerman (ed.) *Handbook of research on writing: history, society, school, individual, text*. London: Routledge, 221–235.
- Beauvais, L., M. Favart, J.-M. Passerault & C. Beauvais 2014. Temporal management of the writing process: effects of genre and organizing constraints in grades 5, 7, and 9. *Written Communication*, 31 (3), 251–279. <https://doi.org/10.1177/0741088314536361>.
- Beers, S. F., T. Mickail, R. Abbott & V. Berninger 2017. Effects of transcription ability and transcription mode on translation: evidence from written compositions, language bursts and pauses when students in grades 4 to 9, with and without persisting dyslexia or dysgraphia, compose by pen or by keyboard. *Journal of Writing Research*, 9 (1), 1–25. <https://doi.org/10.17239/jowr-2017.09.01.01>.
- Bell, A. 1991. *The language of news media*. Oxford: Blackwell.
- Bell, A. & P. Garrett 1998. *Approaches to media discourse*. Oxford: Blackwell.
- Bernstein, J. H. 2015. Transdisciplinarity. A review of its origins, development, and current issues. *Journal of Research Practice*, 11 (1), Article R1. <http://jrp.icaap.org/index.php/jrp/article/view/510/412>.
- Brumfit, C. J. 1995. Teacher professionalism and research. In G. Cook & B. Seidlhofer (eds) *Principle and practice in applied linguistics*. Oxford: Oxford University Press, 27–41.
- Compton-Lilly, C. 2014. The development of writing habitus. A ten-year case study of a young writer. *Written Communication*, 31 (4), 371–403. <https://doi.org/10.1177/0741088314549539>.

- Chen, Y-m. 2009. Quotation as a Key to the Investigation of Ideological Manipulation in News Trans-Editing in the Taiwanese Press. *TTR – traduction, terminologie, rédaction*, 22 (2), 203–238. <https://doi.org/10.7202/044830ar>.
- Cilliers, P. 1998. *Complexity and postmodernism. Understanding complex systems*. London: Routledge.
- Dai, H. & X. Xu 2014. Sexism in news: a comparative study on the portray of female and male politicians in *The New York Times*. *Open Journal of Modern Linguistics*, 4, 709–719. <http://www.scirp.org/journal/ojml>.
- Davis, H. H. 1985. Discourse and media influence. In T. A. Van Dijk (ed.) *Discourse and communication: new approaches to the analysis of mass media discourse and communication*. Berlin: De Gruyter Mouton, 44–59.
- Dempsey, N. P. 2010. Stimulated recall interviews in ethnography. *Qualitative Sociology*, 33, 349–367. <https://doi.org/10.1007/s11133-010-9157-x>.
- DiPardo, A. 1994. Stimulated recall in research on writing: an antidote to “I don’t know, it was fine”. In P. Smagorinsky (ed.) *Speaking about writing – reflections on research methodology*. Thousand Oaks, CA: Sage, 163–181.
- Ekström, M. 2006. Interviewing, quoting, and the development of modern news journalism. A study of the Swedish press 1915–1995. In M. Ekström, Å. Kroon & M. Nylund (eds) *News from the interview society*. Göteborg: Nordicom, 21–48.
- Engeström, Y. 1987. *Learning by expanding: an activity-theoretical approach to developmental work research*. Helsinki: Orienta-Konsultit.
- Fairclough, N. 1992. *Discourse and social change*. Cambridge: Polity Press.
- Fang, I. 1991. *Writing Style Differences in Newspaper, Radio and Television News*. Minnesota, MN: Center for Interdisciplinary Studies of Writing, University of Minnesota.
- Fürer, M. 2018. *Modeling, scaling and sequencing writing phases of Swiss television*. Doctoral dissertation. University of Bern.
- Glaser, B. & Strauss, A. L. 1967. *The discovery of grounded theory: strategies for qualitative research*. London: Wiedenfeld and Nicholson.
- Grésillon, A. 2014. Erfahrungen mit Textgenese. Critique génétique und Interpretation. In W. Lukas, R. Nutt-Kofoth & M. Podewski (eds) *Text – Material – Medium. Zur Relevanz editorischer Dokumentationen für die literaturwissenschaftliche Interpretation*. Berlin: De Gruyter Mouton, 66–79.
- Grésillon, A. & D. Perrin 2014. Methodology: from speaking about writing to tracking text production. In E.-M. Jakobs & D. Perrin (eds) *Handbook of writing and text production*. Berlin: De Gruyter Mouton, 79–111.
- Haapanen, L. 2011. Sitaattien tehtävät ja tekeminen kaunokirjallis-journalistisissa lehtijutuissa. *Media & viestintä*, 34 (3), 64–89.
- *Haapanen, L. 2016a. Haastattelupuheen rekontekstualisointi sitaateiksi lehtijuttuun. *Virittäjä*, 120 (2), 218–254. <https://journal.fi/virittaja/article/view/51313>.
- *Haapanen, L. 2016b. Monologisation as a quoting practice. Obscuring the journalist’s involvement in written journalism. *Journalism Practice*, 11 (7), 820–839. <https://doi.org/10.1080/17512786.2016.1208057>.
- Haapanen, L. 2016c. Siteeraus lehtijutuissa. In L. Tiittula and P. Nuolijärvi (eds) *Puheesta tekstiksi. Puheen kirjallisen esittämisen alueita, keinoja ja rajoja*. SKS:n toimituksia 1424. Helsinki: Suomalaisuuden Kirjallisuuden Seura, 192–224.
- Haapanen, L. 2017a. Quoting in written journalism (doctoral thesis consisting of four articles marked with * in the list of references. Helsinki: University of Helsinki. [http://urn.fi/URN:ISBN:ISBN 978-951-51-2952-9](http://urn.fi/URN:ISBN:ISBN%20978-951-51-2952-9).
- *Haapanen, L. 2017b. Directly from interview to quotations? Quoting practices in written journalism. In R. Laury, M. Etelämäki, and E. Couper-Kuhlen (eds) *Combining Clauses*

- and *Actions in Social Interaction*. Studia Fennica. Helsinki: Suomalaisen Kirjallisuuden Seura, 201–239.
- *Haapanen, L. 2018. Rethinking quoting in written journalism: an intertextual chain from an interview into quotations. In M. Burger (ed.) *Investigating journalism practices: combining media discourse analysis and newsroom ethnography*. Cahier de l'Institut de Linguistique et des Sciences du Langage 54. Lausanne: Université de Lausanne, 95–118.
- Haapanen, L. Forthcoming. Modelling quoting in newswriting. A structure of the process of, and a framework for further studies on, newswriting [article manuscript].
- Hansen, G. 2006. Retrospection methods in translator training and translation research. *Journal of Specialised Translation*, 5, 2–40.
- Helle, M. 2000. Disturbances and contradictions as tools for understanding work in the newsroom. *Scandinavian Journal of Information Systems*, 12 (1–2), 81–113.
- Helle, M. 2010. *Toimitustyö muutoksessa. Toiminnan teoria ja mediakonseptin käsite tutkimuksen ja kehittämisen kehyksenä*. Tampere: University of Tampere. <http://urn.fi/urn:isbn:978-951-44-8313-4>.
- Helle, M. & M. Töyry 2009. Media concept as a tool for analyzing change in media. In P. Oittinen & H. Saarela (eds) *Print Media. Principles, processes and quality*. Helsinki: Paper Engineers' Association/Paperi ja Puu Oy, 497–530.
- Hicks, T. & D. Perrin 2014. Beyond single modes and media: writing as an ongoing multimodal text production. In E.-M. Jakobs & D. Perrin (eds) *Handbook of writing and text production*. Berlin: De Gruyter Mouton, 232–253.
- Hodge, B. 1979. Newspapers and communities. In R. Fowler, B. Hodge, G. Kress & T. Trew (eds) *Language and control*. London: Routledge, 157–174.
- Jakobs, E.-M. 2005. Writing at work. In E.-M. Jakobs, K. Lehnen & K. Schindler (eds) *Schreiben am Arbeitsplatz*. Frankfurt am Main: Verlag für Sozialwissenschaften, 13–40. https://doi.org/10.1007/978-3-322-80777-9_2.
- Jakobs, E.-M. & D. Perrin (eds) 2014. *Handbook of writing and text production*. Berlin: De Gruyter Mouton.
- Jakobsen, A. L. 2003. Effects of think aloud on translation speed, revision and segmentation. In F. Alves (ed.) *Triangulating Translation*. Amsterdam: John Benjamins, 69–95. <https://doi.org/10.1075/btl.45.08jak>.
- Johnson Barella, D. 2005. La literalidad en el uso de las citas directas en las noticias de la prensa regional Navarra. Dos casos: Diario de Noticias y Diario de Navarra. *Comunicación y Sociedad*, 17 (2), 109–140. <http://hdl.handle.net/10171/8249>.
- Lammers, J. C. & V. L. Marsh 2018. "A writer more than . . . a child": a longitudinal study examining adolescent writer identity. *Written Communication*, 35 (1), 89–114. <https://doi.org/10.1177/0741088317735835>.
- Latif, M. M. 2008. A state-of-the-art review of the real-time computer-aided study of the writing process. *International Journal of English Studies*, 8 (1), 29–50.
- Lehrer, A. 1989. Remembering and representing prose: quoted speech as a data source. *Discourse Processes*, 12 (1), 105–125. <https://doi.org/10.1080/01638538909544721>.
- Leijten, M. & L. Van Waes 2013. Keystroke logging in writing research: using inputlog to analyze and visualize writing processes. *Written Communication*, 30 (3), 358–392. <https://doi.org/10.1177/0741088313491692>.
- Leppänen, L., M. Munezero, M. Granroth-Wilding & H. Toivonen 2017. Data-Driven News Generation for Automated Journalism. In *The 10th International Natural Language Generation conference, Proceedings of the Conference*. Stroudsburg: Association for Computational Linguistics, 188–197.
- Lillis, T. 2008. Ethnography as method, methodology, and "deep theorizing": closing the gap

- between text and context in academic writing research. *Written Communication*, 25 (3), 353–388. <https://doi.org/10.1177/0741088308319229>.
- Linell, P. 1998. *Approaching dialogue – talk, interaction and contexts in dialogical perspectives* (IMPACT: Studies in Language and Society). Amsterdam: John Benjamins.
- Lindgren, E. & K. P. H. Sullivan 2003. Stimulated recall as a trigger for increasing noticing and language awareness in the L2 writing classroom: a case study of two young female writers. *Language Awareness*, 12 (3–4), 172–186. <https://doi.org/10.1080/09658410308667075>.
- Lyle, J. 2003. Stimulated recall: a report on its use in naturalistic research. *British Educational Research Journal*, 29, 861–78. <https://doi.org/10.1080/0141192032000137349>.
- Mangen, A., L. G. Anda, G. H. Oxborough & K. Brønnick 2015. Handwriting versus keyboard writing: effect on word recall. *Journal of Writing Research*, 7 (2), 227–247. <https://doi.org/10.17239/jowr-2015.07.02.1>
- Manninen, V. J. E. 2017. Sourcing practices in online journalism: an ethno-graphic study of the formation of trust in and the use of journalistic sources. *Journal of Media Practice*, 18 (2–3), 212–228. <https://doi.org/10.1080/14682753.2017.1375252>
- Matsushita, K. 2016. *Risk management in the decision-making process of English-Japanese news translation*. Doctoral dissertation. Graduate School of Intercultural Communication, Rikkyo University.
- Merminod, G. 2018. *La trajectoire d'une histoire: formes, usages et idéologies du raconter journalistique*. Doctoral dissertation. University of Lausanne.
- Mitchell, A. & T. Rosenstiel 2000. Don't touch that quote. *Columbia Journalism Review*, 38 (5), 34–36.
- Morgan, J. H., C.-Y. Cheng, C. Pike & F. E. Ritter 2013. A design, tests and considerations for improving keystroke and mouse loggers. *Interacting with Computers*, 25 (3), 242–258. <https://doi.org/10.1093/iwc/iws014>.
- Myers, G. 2005. Applied linguists and institutions of opinion. *Applied Linguistics*, 26 (4), 527–544. <https://doi.org/10.1093/applin/ami025>.
- Olive, T. & M.-L. Barbier 2017. Processing time and cognitive effort of longhand note taking when reading and summarizing a structured or linear text. *Written Communication*, 34 (2), 224–246. <https://doi.org/10.1177/0741088317699898>.
- Perrin, D. 1999. Woher die Textbrüche kommen. Der Einfluß des Schreibprozesses auf die Sprache im Gebrauchstext. *Zeitschrift für Deutsche Sprache*, 2, 134–155.
- Perrin, D. 2003. Progression analysis (PA): investigating writing strategies at the workplace. *Journal of Pragmatics*, 35 (6), 907–921. [https://doi.org/10.1016/S0378-2166\(02\)00125-X](https://doi.org/10.1016/S0378-2166(02)00125-X).
- Perrin, D. 2013. *The linguistics of newswriting*. Amsterdam: John Benjamins.
- Perrin, D. Forthcoming. Working with large data corpora in real-life writing research. In K. Sullivan & E. Lindgren (eds) *Observing writing: insights from keystroke logging*. Amsterdam: Brill.
- Pogner, K.-H. 2012. A social perspective on writing in the workplace. Communities of Discourse (DC) and Communities of Practice (CoP). In A. Rothkegel & S. Ruda (eds) *Communication on and via technology*. Berlin: De Gruyter Mouton, 83–107.
- Saldaña, J. 2003. *Longitudinal research: analyzing change through time*. Walnut Creek, CA: AltaMira.
- Saldaña, J. 2009. *The coding manual for qualitative researchers*. Los Angeles, CA: Sage.
- Sarangi, S. 2006. The conditions and consequences of professional discourse studies. In R. Kiely, P. Rea-Dickins, H. Woodfield & G. Clibbon (eds) *Language, culture and identity in applied linguistics*. London: Equinox, 199–220.

- Satoh, A. 2001. Constructing imperial identity: how to quote the imperial family and those who address them in the Japanese press. *Discourse Society*, 12 (2), 169–94. <https://doi.org/10.1177/0957926501012002003>.
- Smagorinsky, P. 1989. The reliability and validity of protocol analysis. *Written Communication*, 6 (4), 463–479. <https://doi.org/10.1177/0741088389006004003>.
- Smagorinsky, P. 1994. Introduction: potential problems and problematic potentials of using talk about writing as data about writing process. In P. Smagorinsky (ed.) *Speaking About Writing: Reflections on Research Methodology*. Thousand Oaks, CA: Sage, 163–181.
- Spelman Miller, K. 2006. Keystroke logging. An introduction. In K. Sullivan & E. Lindgren (eds) *Computer keystroke logging and writing. Methods and applications*. Amsterdam: Elsevier, 1–9.
- Stratman, J. F. & L. Hamp-Lyons 1994. Reactivity in concurrent think-aloud protocols. Issues for research. In P. Smagorinsky (ed.) *Speaking about writing. Reflections on research methodology*. Thousand Oaks: Sage, 89–112.
- Strauss, A. 1987. *Qualitative analysis for social scientists*. Cambridge: Cambridge University Press.
- Strauss, A. & J. Corbin 1998. *Basics of qualitative research: techniques and procedures for developing grounded theory* (2nd edition). Thousand Oaks: Sage.
- Tavory, I. & S. Timmermans 2009. Two cases of ethnography: grounded theory and the extended case method. *Ethnography*, 10 (3), 243–263. <https://doi.org/10.1177/1466138109339042>.
- Töyry, M. 2005. *Varhaiset naistenlehdet ja naisten elämän ristiriidat. Neuvotteluja lukijasopimuksesta*. Doctoral dissertation. Viestinnän julkaisuja 10. Helsinki: Helsingin yliopiston viestinnän laitos.
- Van Hout, T. 2015. Between text and social practice: balancing linguistics and ethnography in journalism studies. In F. Copland, S. Shaw & J. Snell (eds) *Linguistic Ethnography: interdisciplinary explorations*. Palgrave advances in language and linguistics. London: Palgrave Macmillan, 71–89. https://doi.org/10.1057/9781137035035_4.
- Van Hout, T. & F. Macgilchrist 2010. Framing the news: an ethnographic view of business newswriting. *Text & Talk*, 30 (2), 169–191. <https://doi.org/10.1515/text.2010.009>.
- Vandendaele, A., L. De Cuyper & E. Van Praet 2015. Beyond ‘trimming the fat’ – the sub-editing stage of newswriting. *Written Communication*, 32 (4), 368–395. <https://doi.org/10.1177/0741088315599391>.
- Velthuis, O. 2016. The production of a newspaper interview: following the story at the business desk of a Dutch daily. *Journalism*, 17 (7), 899–914. <https://doi.org/10.1177/1464884915583771>.
- Virkkunen, J. 2006. Dilemmas in building shared transformative agency. *@tivités*, 3 (1), 44–66. <https://doi.org/10.4000/activites.1850>.
- Virkkunen, J. 2007. Collaborative development of a new concept for an activity. *@tivités*, 4 (2), 158–164. <https://doi.org/10.4000/activites.1769>.
- Wortham, S. & A. Reyes 2015. *Discourse analysis beyond the speech event*. London: Routledge.
- Zampa, M. & D. Perrin 2016. Arguing with oneself. The writing process as an argumentative soliloquy. *Journal of Argumentation in Context*, 5 (1), 9–28.