

PDF hosted at the Radboud Repository of the Radboud University Nijmegen

The following full text is a preprint version which may differ from the publisher's version.

For additional information about this publication click this link.

<http://hdl.handle.net/2066/158993>

Please be advised that this information was generated on 2021-06-24 and may be subject to change.

1 K. de Haan-Vis and W. Spooren

2 **Informalization in Dutch journalistic** 3 **subgenres over time**

4
5
6
7 **Abstract:** This paper investigates the claim that a trend of informalization can be
8 found in public discourse. This hypothesis is studied in the context of Dutch
9 news reports: recent news reports show a greater amount of subjective elements
10 than older reports. To test the hypothesis, a corpus linguistic analysis of five
11 major Dutch newspapers from 1950/1 and 2002 was carried out, in which the fre-
12 quencies of subjectivity indicators were established in three different journalistic
13 subgenres: hard news, opinion, and background news. On the basis of a pilot
14 study, a distinction was made between reporter's text and quoted text. This
15 proved central to our findings: overall, the more recent news reports do not dis-
16 play a higher degree of informalization; reporter's text even shows a decrease of
17 subjective elements. In contrast, quoted text shows a higher degree of informali-
18 zation across the board. Although subgenre was involved in various interactions
19 with year of publication, the pattern was too irregular to reach firm conclusions
20 about the role of subgenre in the informalization process.

21 **Keywords:** Biber; conversationalization; corpus analysis; diachronic change;
22 Dutch national newspapers; informalization; journalism; journalistic subgenres;
23 logistic regression; personalization; quantitative analysis; quotation; subjectivity,
24 linguistic features of
25

26 27 **1 Introduction¹**

28
29
30 It has been suggested that “a major change in discursive practices affecting
31 many public institutions in contemporary society is the ‘conversationalization’
32 of public discourse [...]” (Fairclough and Wodak 1997, p. 265). This change,
33

34
35 **1** We thank three anonymous reviewers for their valuable comments on a previous version of
36 this paper, and Lieke Verheijen for correcting our English. We thank Onno Huber (VU University
37 Amsterdam) for his assistance in assembling and analysing the corpus, and Gerben Mulder (VU
38 University Amsterdam) for his help in the statistical analyses.

39 **K. de Haan-Vis**, Dedicon, Grave

40 **W. Spooren**, Centre for Language Studies, Radboud University Nijmegen

1 also indicated as ‘informalization’, involves “the modelling of public discourse
2 upon the discursive practices of ordinary life, ‘conversational’ practices in a broad
3 sense” (Fairclough 1994, p. 253). It entails a change in the style and register of
4 many genres, among which journalistic prose, since World War II.

5 Journalistic prose consists of many subgenres, such as commentaries, news
6 reports, and interviews. In this chapter, we take genres and subgenres to be
7 experience-based, conventionalized conceptualizations of fixed combinations
8 of language or code, context and text, much along the lines of Steen (2011).
9 We regard genre knowledge as schematic knowledge: becoming experienced
10 in a discourse culture means learning the combinations of language, goals
11 and forms of that culture, to perform certain communicative tasks. Typically, a
12 western news report brings a more or less objective report of what happened,
13 how it happened, who was involved, and what the consequences were (context).
14 Such a report is typically expressed in fairly objective language, with a strong
15 focus on the ideational function (Halliday, 1974) (code). The typical discourse
16 form to express such a news event has an inverted structure, with the main
17 news occurring first, followed by additional details and background information
18 (text; Knobloch et al. 2004). Knowledge of the genre combinations helps us to
19 process information quickly and efficiently. Such knowledge is typically culture
20 dependent: for example, Iranian journalists frequently use chronological order-
21 ing in delivering news reports (Rafiee, in preparation).

22 This chapter focuses on genre change at the level of the language, or code.
23 From a linguistic viewpoint, what the changes toward informalization have in
24 common is a shift in attention from the object, i.e. the news, to the subject, i.e.
25 the journalist and sources he or she quotes in the news report. In other words,
26 the informalization hypothesis presumes that the journalist and the news sources
27 have become more central as speakers. This centrality of speakers is linguistically
28 reflected in the expression of self: the perspective of the writer or another subject
29 of consciousness mentioned in the text. These are all aspects of what in linguistic
30 theorizing has become known as subjectivity (Lyons 1982; Langacker 1990;
31 Traugott 1989). Note that studying the informalization hypothesis in terms of
32 changes in expression of linguistic subjectivity is by no means the only way to
33 study the process of informalization. Others have looked at, for example, changes
34 in lexis (such as the use of metaphors; Pasma 2011) or grammatical developments
35 (e.g. the use of contractions and questions; Pearce 2005). Our focus is on subjec-
36 tivity because it is a linguistic phenomenon that is particularly apt at reflecting
37 changes with respect to informalization: as will be discussed below, there seems
38 to be a strong link between informal and subjective language.

39 Hundt and Mair (1999) described journalistic prose as an ‘agile’ genre com-
40 pared to academic prose, in that journalistic prose is more open to innovations

1 such as ‘colloquialization,’ as observed in increases of first and second person
 2 pronouns, contractions, etc. Since informalization would imply an increase over
 3 time of “conversation” between two voices, the journalist and his/her reader, a
 4 crucial question is whether these voices have become more prominent over time.

5 Another question is whether the process of informalization is homogeneous
 6 across subgenres. Journalistic texts can be written in various subgenres; some of
 7 them are generally considered to be objective reports of news events, whereas
 8 others allow for an opinionated voice from the journalist (news commentaries,
 9 for example). An in-between subgenre is the background article, which typically
 10 expands on news reports that were published earlier, and allows the writer more
 11 freedom.

12 A journalistic text is multivoiced: when describing a news event, a journalist
 13 often quotes sources for reasons of vividness (quotes are often seen as a means
 14 of enlivening a text) and distancing (the journalist may not want to take respon-
 15 sibility for a particular piece of information and may therefore attribute it to a
 16 specific source). To investigate the multivoicedness of journalistic texts, a dis-
 17 tinction was made in this paper between quoted text and journalist’s text (or:
 18 ‘reporter’s text’).

19 The linguistic analysis of journalistic subgenres over time is a particularly
 20 interesting case in point for the present volume’s goal for several reasons. Firstly,
 21 such a diachronic investigation of changes in linguistic features presents a
 22 view on ‘genres-in-development’: how is genre change expressed linguistically?
 23 Secondly, the scale of the analysis, in terms of both corpus size and the number
 24 of linguistic features analyzed, plus the rigid statistical comparison of the
 25 patterns over time, newspapers and subgenres, makes this an enterprise that
 26 could be a model for future linguistic analyses of genre change. Hence, our
 27 research also has strong methodological value.

30 2 Theoretical background

32 For the analysis of subjective elements, we followed a very pragmatic approach
 33 by selecting operationalizations mentioned in the literature (see also Vis, Sanders
 34 and Spooren 2012). More specifically, we selected features described by Bekker
 35 (2006), Biber (2004), Scheibman (2002), Wiebe (1994) and Wiebe et al. (2005).
 36 The list was completed by the descriptions of the grammatical categories of
 37 stance and modality in the grammars of Biber et al. (1999) and Haeseryn et al.
 38 (1997). The appendix gives a detailed operationalization of our subjectivity lexicon.
 39 Here we list the general categories.

- 1 1. First and second person pronouns explicitly refer to the writer or the
2 addressee, or, in case of a quotation, to a character mentioned in the text.
3 That is why many authors take them as clear indicators of subjectivity (a.o.
4 Scheibman 2002).
- 5 2. Modal elements (such as adverbs [*possibly*], verbs [*must*] and particles
6 [*only*]) reflect the writer's attitude toward what is expressed in the text. As
7 a consequence, they have been mentioned as indicators of subjectivity (a.o.
8 Biber, Johansson, Leech, Conrad, and Finegan 1999; Haeseryn et al. 1997;
9 Precht 2000).
- 10 3. Intensifiers or degree adverbs express a writer's estimate, on an intensity
11 scale, of the degree to which an utterance is true (Bolinger 1972) and are
12 often tied up to evaluations by a subject of consciousness (Thompson and
13 Hunston 2000).
- 14 4. Cognitive verbs like *know*, *doubt* and *answer* "can be interpreted as express-
15 ing author/speaker attitude and assessment of likelihood" (Biber 2004). As a
16 result, they reflect a subject's point of view.
- 17 5. Exclamations and questions also reflect a subject's point of view, as they
18 suggest that a subject expresses his/her mood or addresses a communica-
19 tional partner. As such, they have been suggested as elements reflecting
20 subjectivity (e.g. Wiebe 1994).
- 21 6. Deixis positions an utterance with respect to the here and now of the current
22 writer/speaker. As a consequence, it reflects the presence of a subject of con-
23 sciousness, which is why it has been mentioned as reflecting epistemological
24 stance (e.g. Bekker 2006).

25
26 Several of these lexicogrammatical indicators of subjectivity have been included
27 in one or more diachronic studies. This section discusses the most prominent
28 of those studies. As will be seen, many indicators show an increase over time,
29 suggesting a shift in written discourse toward more conversational and informal
30 norms. Table 1 presents a summary of these findings.

31 A number of these studies focus on news genres. These are discussed in the
32 remainder of this section. Westin and Geisler (2002) used Biber's (1988) multi-
33 feature/multi-dimensional model and the concomitant characteristic features
34 for a study of informalization in British newspaper discourse. In an analysis of
35 554 editorials from three up-market British newspapers published between 1900
36 and 1993, they found that the editorials became less abstract (dimension 5) and
37 less dependent on referential elaboration (dimension 3, marked by such features
38 as an increase of temporal and spatial adverbials), which resulted in more
39 informal language. Similar to Biber and Finegan's (2001) findings for news
40

Table 1: Summary of findings in previous diachronological studies

	↑	↓	↔
1st, 2nd person pronouns	Hundt and Mair (1999); Leech et al. (2009); Pearce (2005); Steen (2003)	Leech et al. (2009); Westin and Geisler (2002)	Westin and Geisler (2002)
modal adverbs	Biber (2004) (written) Pearce (2005) (election broadcasts)	Biber (2004) (news, medical)	
intensifiers			Pearce (2005); Westin and Geisler (2002)
modals	Pearce (2005); Steen (2003)	Biber (2004); Leech et al. (2009)	Westin and Geisler (2002)
semi-modals	Biber (2004); Leech et al. (2009)		
cognitive verbs	Biber (2004); Pearce (2005)		Westin and Geisler (2002)
exclamation marks		Leech et al. (2009)	
questions and question marks	Leech et al. (2009); Westin and Geisler (2002)	Leech et al. (2009)	Westin and Geisler (2002)

Note: ‘↑’: increase; ‘↓’: decrease; ‘↔’: no change; studies of news reports are printed in boldface.

reportage, no regular pattern was observed with respect to dimension 1 (distinguishing between involved and informational discourse, marked by such features as private verbs, first and second person pronouns, hedges, amplifiers, ‘wh’-questions, and possibility modals). Westin and Geisler’s analyses indicate that the changes mainly took place during the latter part of the twentieth century, since the last period included in their study – representing the years 1960 up to 1993 – differs from the preceding two periods on three of the dimensions.

Steen (2003) provided a study similar to Westin and Geisler’s: he, too, used Biber’s (1988) multi-feature/multi-dimensional model to study the stylistic changes in editorials from the British press, in this case *The Times*. The period of publication of the 120 editorials (85,000 words) overlaps with the last period from Westin and Geisler’s study: 1950 to 2000. Steen investigated four styles, labeled as “involved”, “informational”, “narrative” and “persuasive”, representing Biber’s dimension 1, 2, and 4. In an analysis of a limited number of features

1 per style (including the subjectivity indicators first and second person pronouns
 2 for the “involved” style, and modal and semi-modal verbs for the “persuasive”
 3 style), Steen found an increase in involvement and persuasion and a decrease in
 4 what he, following Biber (1988), marks as narrativity: past tense, perfect aspect
 5 and third person. This can be interpreted as a stylistic change toward conversa-
 6 tions, as these are typically involved and persuasive in nature and show few
 7 markings of narrativity.

8 Similar to Biber and Finegan (2001), Hundt and Mair (1999) also observed a
 9 difference in the shift toward a more involved and oral style between several
 10 written genres. In a study of news and academic prose (320,000 words) from
 11 the Brown corpora, Hundt and Mair (1999) found evidence of a “colloquializa-
 12 tion” of the norms of written English, that is, a narrowing of the gap between
 13 spoken and written norms. However, they also observed that the two genres
 14 differ in their openness to innovation: news is relatively receptive to innovation
 15 (“agile”), whereas academic prose is more reluctant to change (“uptight”). This
 16 openness to innovation in journalistic prose is shown by an increase in the use
 17 of first and second person pronouns, contractions, sentence-initial conjunctions,
 18 phrasal verbs, and progressive aspect.

19 Using the same corpora, Leech et al. (2009) described the diachronic change
 20 of a wide range of grammatical topics, including modal and semi-modal verbs,
 21 progressives, subjunctives, passives, genitives, and relative clauses in con-
 22 temporary written English. Leech et al. (2009) found evidence for colloquializa-
 23 tion in several features, such as an increase of contractions, ‘not’-negation versus
 24 ‘no’-negation, progressives, ‘get’-passives, a decrease of ‘be’-passives, a decline
 25 of ‘wh’-relative clauses and an increase of ‘that’- and zero relative clauses. There
 26 was also an increase of first and second person pronouns in the American
 27 English corpora, an increase of questions in the corpus overall (although they
 28 decreased in the press section), an increase in the use of a representative set of
 29 so-called semi-modals (e.g. ‘be going to’, ‘(had) better’, ‘need to’), and a sub-
 30 stantial decline in the frequency of so-called core modals (e.g. ‘will’, ‘would’,
 31 ‘must’, ‘might’). However, Leech et al. (2009: 245) observed that colloquializa-
 32 tion does not occur in all cases: for example, they were surprised to find that
 33 the use of exclamation marks has declined over time in both American and
 34 British English, and that the use of first person pronouns in British English
 35 “shows a mysterious anti-colloquial trend” (Leech et al. 2009: 247).

36 Biber (2004) observed that in news, drama, personal letters, and medical
 37 prose speakers and writers have become more inclined to express their opinions,
 38 attitudes and emotions, especially over the last fifty years of the study. In an
 39 analysis of 560 texts from the ARCHER corpus from a period from 1650 to 1990,
 40

1 he investigated three major structural types of stance marking from Biber et al.
2 (1999): modal and semi-modal verbs, stance adverbials (e.g. ‘hopefully’, ‘un-
3 doubtedly’, ‘evidently’), and stance complement clause constructions (e.g. ‘suggest
4 that’, ‘expect that’, ‘possible that’, ‘hope to’, ‘believe to’, ‘easy to’). The results
5 differed between the four genres: modal verbs declined in all genres, whereas
6 semi-modal verbs showed an increase, but this was restricted primarily to drama
7 and personal letters. Stance adverbials increased steadily in use across all
8 genres in the nineteenth century, but in the most recent period this increase
9 did not continue in news and medical prose. Finally, the complement clause
10 constructions also underwent a steady increase in frequency across the periods,
11 the shift toward greater use of stance verb + complement being especially
12 characteristic of drama and personal letters, and the stance adjective/noun +
13 complement of newspapers and to a lesser extent medical prose.

14 In sum, the discussion above shows that there are several lexicogrammatical
15 indicators of subjectivity that have been included in one or more diachronic
16 studies. Many of the indicators have seen an increase over time, which suggests
17 a shift in written discourse toward more conversational and informal norms.

18 None of the studies mentioned have investigated the informalization hypoth-
19 esis for Dutch, and none have distinguished between subgenres in newspaper
20 texts. Our present study aims to fill that gap.

23 3 Method

25 3.1 Corpus selection

27 To answer our research questions, a diachronic corpus of Dutch written news
28 data was compiled, representing different subgenres and different time periods.
29 News reports were selected from five Dutch national newspapers from the periods
30 1950/1951 and 2002. This interval reflects two pivotal instances in societal develop-
31 ment: the first point is directly after World War II and before the presumed inform-
32 alization development, while the second point is after the presumed informaliza-
33 tion. The year 2002 was chosen because it fell between two relevant developments
34 within journalism: after a period of professionalization, but before the broad
35 spread of digitization of journalism. 2002 was also a relevant year because in
36 a different study, the data from the news corpus have been compared with
37 those from conversation data in the so-called Corpus of Spoken Dutch, which
38 was collected around 2002 (see Vis 2011).

1 The newspapers selected for the corpus were national newspapers published
2 in both periods of investigation. Our choice of newspapers reflects the distinction
3 between quality newspapers and more tabloid-like newspapers (although the
4 difference is much smaller in the Netherlands compared to, for example, the
5 United Kingdom; Deuze 2005). *NRC Handelsblad*, *Volkscrant* and *Trouw* are
6 generally considered quality newspapers, whereas *De Telegraaf* and *Algemeen*
7 *Dagblad* are more tabloid-like.

8 The data for 1950/1 were collected from prints in the libraries of VU University
9 Amsterdam and the University of Amsterdam. The data for one newspaper
10 (*Algemeen Dagblad*) came from high-quality scans made available by the
11 National Library (Koninklijke Bibliotheek). The articles were selected semi-
12 randomly: for each newspaper, three weeks were chosen; all editions of that
13 week were included in the corpus until a total of 50 pages was reached. The
14 weeks were more or less evenly distributed over the year. If a page contained
15 too many advertisements, it was disregarded and a new page was selected.
16 The pages were digitized using OCR. Post-processing was applied to eliminate
17 recognition errors.

18 The data for 2002 were collected from the LexisNexis database. For each
19 newspaper, a week was selected semi-randomly. From each specific week, all
20 articles of that week were included in the corpus. The weeks were more or less
21 evenly distributed over the year.

22 The articles selected from the corpus have been aggregated into three sub-
23 genres: hard news, opinion news and background news (following Bell 1991).
24 Hard news encompasses front page news, domestic news, foreign news, and
25 business news; opinion news comprises editorials, opinion news, and arts
26 (reviews); background news contains the categories sports, and science, and
27 the section on 'showbiz' news in *De Telegraaf* from 2002 and the articles in the
28 1950s newspapers categorized as miscellaneous.

29 Both corpora were annotated automatically for part of speech and lemma
30 information using the Tadpole tagger (Van den Bosch et al., 2007). During this
31 phase, the texts were also transformed from a text format to an XML format.
32 Because of its unique character, the resulting corpus had been made available
33 through a so-called Clarin action as the VU-DNC corpus (cf. <http://tst-centrale.org/nl/producten/corpora/vu-dnc-corpus/6-79>).

35 The corpus consists of approximately 1,900,000 words in 6,618 news reports
36 from three different subgenres (hard news, opinion news, background news).
37 Table 2 presents an overview of the corpus.

38
39
40

1 **Table 2:** Corpus size (in words)

	Hard news	Opinion	Background	Total	
1950/1951					
5	<i>Algemeen Dagblad</i>	90,247	34,798	60,790	185,835
6	<i>NRC (Handelsblad)</i>	117,574	36,375	31,041	184,990
7	<i>De Telegraaf</i>	105,811	30,516	42,287	178,614
8	<i>Trouw</i>	135,761	13,484	32,410	181,655
9	<i>De Volkskrant</i>	115,648	23,954	60,878	200,480
9	Total	565,041	139,127	227,406	931,574
2002					
11	<i>Algemeen Dagblad</i>	117,083	28,046	75,951	221,080
12	<i>NRC (Handelsblad)</i>	164,416	53,846	41,767	260,029
13	<i>De Telegraaf</i>	76,658	20,755	34,700	132,113
14	<i>Trouw</i>	86,270	40,551	22,570	149,391
15	<i>De Volkskrant</i>	114,484	46,687	47,275	208,446
16	Total	558,911	189,885	222,263	971,059

18 3.2 Subjectivity analysis

20 An automatic analysis was made of indicators of subjectivity at the lexicogram-
 21 matical level. Drawing on previous research on subjectivity and stance, a list of
 22 lexicogrammatical indicators of subjectivity was compiled. Examples are lin-
 23 guistic features such as first and second person pronouns, modal elements
 24 (verbs, adverbs, and particles) and intensifiers. An overview of the list of features
 25 used for this analysis can be found in the Appendix.

26 On the basis of the literature and previous analyses, a distinction was made
 27 between lexicogrammatical indicators that reflected reporter subjectivity (found
 28 in the reporter's words) and lexicogrammatical indicators that reflected source
 29 subjectivity (found in direct quotations).

32 3.3 Identifying quoted words

34 In order to be able to distinguish between the reporter's text and the quoted
 35 text, the news reports were marked for direct quotations on the basis of quota-
 36 tion marks. The choice for direct quotation over, for example, indirect and free
 37 indirect discourse was made deliberately: the quotation marks are an explicit
 38 indication that the journalist is not responsible for the exact wording of the
 39 quote, so that responsibility, including the choice for subjective elements, lies
 40 with the quoted source. We realize that relying on quotation marks is a fairly

1 rough measure to indentify direct quotations. However, the size of the corpus
2 did not allow us to identify the quotations manually. To ensure that this method
3 was reliable, samples of quotation annotations were checked manually. In a
4 total of 70,000 words, only two percent was annotated incorrectly. This was
5 considered sufficient for present purposes.

8 3.4 Statistical analysis

9
10 Three subsets of materials were distinguished: the complete texts (the text
11 “overall”), the reporter’s text, and the direct quotations. For each of the three
12 subsets and for each of the subjectivity markers, a logit analysis was carried
13 out, in which for each text the number of observed occurrences of subjectivity
14 markers was related to the number of words in the text. A logit analysis is a
15 type of logistic regression in which the probability of frequencies is calculated
16 using the logit function: the logit of the number of words $\text{logit}(F)$ is defined as
17 $\text{log}(F/(N-F))$, where F is the frequency of occurrence of the subjectivity markers
18 and N is the number of words in the text/reporter’s text/direct quotations in the
19 text. A logit analysis was chosen because it can deal with samples that differ
20 greatly in the frequency of the phenomenon under investigation. Moreover, such
21 an analysis allows one to study interactions between predictors. In that sense, it
22 resembles an analysis of variance.

23 In the “overall” analysis, the number of occurrences was related to the total
24 number of words in the text; in the analysis of reporter’s words, it was related
25 to the total number of reporter words; and in the analysis of quoted words,
26 it was related to the total number of quoted words. The predictor variables in
27 the analysis were Year (1950/1951, 2002), Newspaper (*Algemeen Dagblad*, *NRC*
28 (*Handelsblad*), *De Telegraaf*, *Trouw*, and *De Volkskrant*) and Subgenre (hard
29 news, opinion, background news). Texts were treated as a random factor. For
30 the analysis the MLwiN software (version 2.18) was used.

31 The analysis was carried out using backward elimination. Only those terms
32 were kept that contributed significantly to the model. If the main effect of Year
33 was in the final model, a separate χ^2 -analysis of Year was carried out.

34 In the following section, we give an overall overview of the results, one for
35 each subset of materials (overall, reporter’s text, quoted text). Space does not
36 allow us to give all results, so we restrict ourselves to findings in which the
37 factor Year is involved. That implies that main effects of Subgenre and News-
38 paper and interactions between Newspaper and Subgenre were disregarded.
39 For a complete overview of the results and statistical tests, we refer to Vis
40 (2011), especially chapter 5 and appendix B.

4 Results

4.1 ‘Overall’ analysis

Table 3 gives a summary of the results for the overall texts. An increase from 1950/1951 to 2002 is indicated by an upward arrow (‘↑’), a decrease is indicated by a downward arrow (‘↓’), and a lack of change is indicated by a horizontal bidirectional arrow (‘↔’). Arrows that are underlined indicate statistical significance. For each of the indicators, the effects in which the year of publication is involved are listed; the corresponding development is displayed in the adjacent column.

For example, for first person singular pronouns, there is a main effect of year (Y) which shows an increase (the arrow in the adjacent column), whereas for second person singular pronouns, there are two two-way interactions in which year of publication is involved: one interaction with newspaper (N) and one with Subgenre (S). The corresponding developments are listed in alphabetical order of the name of the newspaper or subgenre: for newspaper, this is *Algemeen Dagblad*, *NRC (Handelsblad)*, *De Telegraaf*, *Trouw*, and *De Volkskrant*, so the first arrow in the column next to Y*N for second person singular pronouns is for *Algemeen Dagblad* (a ‘↑’ for increase), the second arrow for *NRC (Handelsblad)* (another ‘↑’), and so on. For the subgenres, the order is ‘hard news’, ‘opinion news’, ‘background news’; hence, for the interaction between year and subgenre the first arrow is for hard news (also a ‘↑’). In case of a three-way interaction, there are three times five arrows; the first set of five is for the five newspapers in alphabetical order in hard news (‘h’), the second set of five for the newspapers in opinion news (‘o’), and the final set for the newspapers in background news (‘b’). The arrows are not based on significance, but rather on direction of the development.

The results of the subjective indicators in the text “overall” can be summarized as follows. For all elements, the year of publication has a significant effect, either as a main effect or in interaction with newspaper and/or subgenre. The only exception is the exclamation mark, which is not used differently over time. The newspapers are mutually divergent in how they use exclamation marks, as are the subgenres, yet these differences do not change over time.

The overview in Table 3 demonstrates that many elements show increases in all (or nearly all) newspapers and subgenres, i.e. first person singular pronouns, second person singular pronouns, cognitive verbs, intensifiers, and modal particles. For the first person singular pronouns, there is a main effect of year; for the second person singular pronouns, the frequency in all newspapers and subgenres is higher in 2002 than in 1950/1; cognitive verbs are used more in 2002

1 **Table 3:** Overview of developments of the indicators of subjectivity in the texts, overall

Indicator	Rel. freq. 1950/1	Rel. freq. 2002	Effect	Development	
1. Pronoun 1p. sg.	19.92	61.29	Y	<u>↑</u>	
2. Pronoun 2p. sg.	10.37	29.91	YxN	<u>↑↑↑↑↑</u>	
			YxS	<u>↑↑↑</u>	
3. Pronoun 1p. pl.	31.25	28.49	YxN	<u>↓↓↓↓↑</u>	
			YxS	<u>↑↓↓</u>	
4. Deictic element	41.18	40.21	YxNxS	h	<u>↓↑↓↓↓</u>
				o	<u>↓↑↓↓↓</u>
				b	<u>↔↑↑↓↔</u>
5. Cognitive verb	87.66	111.76	YxNxS	h	<u>↑↑↑↑↑</u>
				o	<u>↑↑↑↓↑</u>
				b	<u>↑↑↑↓↑</u>
6. Exclamation mark	4.20	3.49	–	↔	
7. Question mark	10.11	14.36	YxNxS	h	<u>↑↑↓↑↑</u>
				o	<u>↑↑↑↑↔</u>
				b	<u>↔↑↓↓↑</u>
8. Intensifier	49.02	65.97	YxN	<u>↑↑↑↑↑</u>	
			YxS	<u>↑↑↑</u>	
9. Modal adverb	25.11	24.32	YxN	<u>↓↓↑↑↑</u>	
			YxS	<u>↓↓↑</u>	
10. Modal particle	98.35	143.83	YxNxS	h	<u>↑↑↑↑↑</u>
				o	<u>↑↑↑↑↓</u>
				b	<u>↑↑↑↑↑</u>
11. Modal verb	169.06	165.85	YxNxR	h	<u>↓↓↔↑↑</u>
				o	<u>↓↓↑↓↓</u>
				b	<u>↑↑↑↓↑</u>

Y = year of publication, N = newspaper, S = subgenre; h = hard news, o = opinion news, b = background news; ↑ = increase, ↓ = decrease, ↔ = no effect of year, so no change.
 Significant changes (p < .05) are underlined. Relative frequency: frequency per 10,000 words.
 Note that these relative frequencies generalize over newspapers and subgenres.

by all newspapers in hard news, by one newspaper in opinion, and by all but one newspaper in background news; all newspapers and all subgenres contain more intensifiers in 2002; and modal particles are used more in 2002 in all newspapers in all subgenres. These increases are in line with the hypothesis that newspapers have undergone a process of informalization.

1 There are no elements for which all newspapers and subgenres uniformly
 2 decrease their use over time. The most convincing case of decrease are modal
 3 verbs: in hard news, two newspapers use significantly fewer modal verbs in 2002
 4 than in 1950/1, and in opinion news, even more newspapers show a decrease in
 5 the number of modal verbs used. However, in background news, four out of five
 6 newspaper show an increase in their use of modal verbs.

7 The remaining elements show an even less straightforward pattern: the inci-
 8 dence of first person plural pronouns is higher in 2002 than in 1950/1 in some
 9 newspapers, whereas it is higher in 1950/1 in others. The same is true for the
 10 three subgenres: in one subgenre, more first person plural pronouns are used
 11 in 2002 than in 1950/1, whereas in the other two subgenres, more are used in
 12 1950/1. Similarly, the newspapers show no uniform pattern for deictic elements:
 13 in both hard news and background news, some newspapers increase their use of
 14 deictic elements over time, whereas others decrease or show no change at all.
 15 The incidence of question marks increases in most newspapers in hard news
 16 and opinion news, but the results are more ambiguous in background news: in
 17 two newspapers, the incidence increases, in two others, it decreases, and in one,
 18 it remains unchanged. Finally, the modal adverbs show increases and decreases
 19 both in newspapers and in subgenres. These ambiguous results do not meet the
 20 expectations that the indicators of subjectivity would show an increase over
 21 time.

22 In sum, it seems that in the text “overall”, some indicators provide evidence
 23 in favour of informalization, whereas other indicators reveal mixed results: only
 24 some newspapers and subgenres show informalization and others do not.

27 4.2 Analysis of reporter’s text

28 In the following two sections, a distinction is made between words from the
 29 journalist/reporter and words from news sources. This section reports the results
 30 for reporter’s text. Table 4 summarizes the results.

31 The results for the indicators in the words of the reporter show that all ele-
 32 ments change in frequency, either increasing or decreasing over time, with the
 33 exception of first person singular pronouns. In contrast with the changes in the
 34 analysis “overall”, all changes in the words of the reporter are rather uniform
 35 across newspapers and subgenres.

36 Five indicators show a rather uniform increase over time for (almost) all
 37 newspapers and subgenres: second person singular pronouns, cognitive verbs,
 38 question marks, intensifiers and modal particles. The previous section has shown
 39 that these indicators also increase in the text “overall”. In the reporter’s text, the
 40

1 **Table 4:** Overview of developments of the 11 indicators in the reporter's text

Indicator	Rel. freq. 1950/1	Rel. freq. 2002	Effect	Development
1. Pronoun 1p. sg.	13.89	14.50	–	↔
2. Pronoun 2p. sg.	7.58	13.57	YxN YxS	<u>↑</u> <u>↑</u> <u>↓</u> <u>↑</u> <u>↑</u> <u>↑</u> <u>↑</u> <u>↓</u>
3. Pronoun 1p. pl.	27.05	9.12	YxNxS	h <u>↓</u> <u>↓</u> <u>↓</u> <u>↓</u> <u>↑</u> o <u>↓</u> <u>↓</u> <u>↓</u> <u>↓</u> <u>↓</u> b <u>↓</u> <u>↓</u> <u>↓</u> <u>↓</u> <u>↓</u>
4. Deictic element	41.32	37.15	YxNxS	h <u>↓</u> <u>↓</u> <u>↑</u> <u>↓</u> <u>↓</u> o <u>↓</u> <u>↓</u> <u>↑</u> <u>↓</u> <u>↓</u> b <u>↓</u> <u>↑</u> <u>↑</u> <u>↓</u> <u>↓</u>
5. Cognitive verb	85.95	102.45	YxNxS	h <u>↑</u> <u>↑</u> <u>↑</u> <u>↑</u> <u>↑</u> o ↔ ↔ <u>↑</u> <u>↓</u> ↔ ↔ b <u>↑</u> <u>↑</u> <u>↑</u> <u>↓</u> <u>↑</u>
6. Exclamation mark	3.53	2.04	Y	<u>↓</u>
7. Question mark	8.94	11.57	YxNxS	h <u>↑</u> <u>↑</u> <u>↓</u> <u>↑</u> <u>↑</u> o <u>↑</u> <u>↑</u> <u>↑</u> <u>↑</u> <u>↓</u> b <u>↓</u> <u>↑</u> <u>↓</u> <u>↓</u> <u>↑</u>
8. Intensifier	48.03	59.84	YxN YxS	<u>↑</u> <u>↑</u> <u>↑</u> <u>↑</u> <u>↑</u> <u>↑</u> <u>↑</u> <u>↑</u>
9. Modal adverb	24.48	19.99	YxN	<u>↓</u> <u>↓</u> <u>↔</u> <u>↓</u> <u>↓</u> <u>↓</u>
10. Modal particle	97.06	131.81	YxNxS	h <u>↑</u> <u>↑</u> <u>↑</u> <u>↑</u> <u>↑</u> o <u>↑</u> <u>↑</u> <u>↑</u> <u>↑</u> <u>↓</u> b <u>↑</u> <u>↑</u> <u>↑</u> <u>↑</u> <u>↑</u>
11. Modal verb	165.86	148.72	YxN YxS	<u>↓</u> <u>↓</u> <u>↓</u> <u>↓</u> <u>↓</u> <u>↓</u> <u>↓</u> <u>↓</u>

Y = year of publication, N = newspaper, S = subgenre; h = hard news, o = opinion news, b = background news; ↑ = increase, ↓ = decrease, ↔ = no effect of year, so no change. Significant changes ($p < .05$) are underlined. Note that these relative frequencies generalize over newspapers and subgenres.

second person singular pronouns are used more in 2002 than in 1950/1 by all newspapers but *De Telegraaf* and in all subgenres except for background news. Similarly, the incidence of cognitive verbs increases in all newspapers in hard and background news (except for *Trouw* and *De Telegraaf*, in the latter subgenre), although the changes in opinion news are slightly more ambiguous.

1 Question marks also increase in all but one newspaper, in hard news and opinion
 2 news. Furthermore, in all newspapers and subgenres, intensifiers are used more
 3 in 2002 than in 1950/1. Finally, the incidence of modal particles increases uni-
 4 formly across all newspapers in all subgenres. These increases correspond to
 5 the informalization hypothesis. It seems that over the course of the second half
 6 of the twentieth century, reporters have become more inclined to express inter-
 7 action with the addressee (in second person singular pronouns and questions)
 8 and interpretation through opinions, attitudes and beliefs through some of the
 9 subjectivity indicators (cognitive verbs, intensifiers, and modal particles).

10 However, contrary to the analysis of the text “overall”, which showed a
 11 mixed result for the remaining indicators, there is a uniform decrease between
 12 1950/1 and 2002 in five elements in the reporter’s text: first person plural pro-
 13 nouns, deictic elements, exclamation marks, modal adverbs, and modal verbs.
 14 First person plural pronouns decrease uniformly in all newspapers in all sub-
 15 genres, although the rates of decrease are not the same. For deictic elements,
 16 the decrease is not as consistent: in hard news, all but two newspapers (*NRC*
 17 (*Handelsblad*) and *Trouw*) use less deictic elements over time, in opinion news,
 18 this is the case in all newspapers but *De Telegraaf*, and in background news, in
 19 all but two newspapers (*NRC(Handelsblad)* and *De Telegraaf*) the use of deictic
 20 elements decreases. For exclamation marks, there is a general pattern of decrease,
 21 which does not differ between the newspapers or subgenres. Modal adverbs
 22 show a decline in frequency in all newspaper but *De Telegraaf*, which does not
 23 change its use over time. Finally, modal verbs occur less in 2002 than in 1950/1
 24 in all newspapers and all subgenres. The only element which does not manifest
 25 a change in frequency of occurrence is the first person singular pronoun.

26 The decreases found in this analysis for first person plural pronouns, deictic
 27 elements, and modal adverbs and the stability over time of first person singular
 28 pronouns were unexpected on the basis of the previous diachronic studies.
 29 However, those studies did not examine the reporter’s text separately. Our re-
 30 sults suggest that, contrary to the informalization hypothesis, reporters have
 31 become more reluctant to refer to themselves and to the here-and-now. It is
 32 possible that this development reflects the professionalization process in Dutch
 33 journalism (Van Vree et al. 2003). Students of journalistic programs are explicitly
 34 trained to use an objective style.

36 4.3 Analysis of direct quotations

37
 38 The final set of results relates to the parts of the texts that come from quoted
 39 news sources, which were presented as direct quotations. In a separate paper,
 40 we have discussed the development of quoted discourse over time in more detail

1 (Vis, Sanders and Spooren 2012). There it is argued that the relative amount of
 2 quoted news increases over time (1950/1: 1153 words per 10,000 words; 2002:
 3 2101/10,000 words). Table 5 gives a summary of the results.

4
 5 **Table 5:** Overview of developments of the 11 indicators in direct quotations

6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26
	Indicator	Rel. freq. 1950/1	Rel. freq. 2002	Effect																
8	1. Pronoun 1p. sg.	78.18	287.91	YxN																
9				YxS																
10	2. Pronoun 2p. sg.	37.31	100.78	YxN																
11	3. Pronoun 1p. pl.	71.77	130.57	YxNxS	h															
12					o															
13					b															
14	4. Deictic element	39.83	55.05	Y																
15	5. Cognitive verb	104.16	156.86	Y																
16	6. Exclamation mark	10.64	10.53	YxN																
17				YxS																
18	7. Question mark	21.40	27.86	YxN																
19	8. Intensifier	58.60	95.67	Y																
20	9. Modal adverb	31.13	45.31	YxS																
21	10. Modal particle	110.80	196.21	Y																
22	11. Modal verb	199.96	248.80	YxNxS	h															
23					o															
24					b															

27 Y = year of publication, N = newspaper, S = subgenre; h = hard news, o = opinion news,
 28 b = background news; ↑ = increase, ↓ = decrease, ↔ = no effect of year, so no change.
 29 Significant changes (p < .05) are underlined. Note that these relative frequencies generalize
 30 over newspapers and subgenres.

31
 32 With respect to the research question concerning the development over time, the
 33 results of the analysis of subjective elements within quotations are relatively
 34 straightforward. For all elements, the year of publication has a significant effect
 35 on the frequency of occurrence, and this effect is always one of increase over
 36 time, except for the exclamation mark, for which the result is more ambiguous.
 37 For most elements, the increase is clear and uniform across all newspapers and
 38 subgenres; only for a few elements, the increase is not as consistent. Table 5
 39 gives an overview of the developments for the 11 indicators.

40

1 Table 5 demonstrates that eight indicators show an increase for (almost)
 2 all newspapers and subgenres: first person singular pronouns, second person
 3 singular pronouns, first person plural pronouns (with the exception of *Algemeen*
 4 *Dagblad* and *Trouw* in opinion news), deictic elements, cognitive verbs, intensi-
 5 fiers, modal particles, and modal verbs (again with the exception of *Trouw* in
 6 opinion news). Only three elements display more variation in their changes
 7 over time: exclamation marks, question marks, and modal adverbs. The use of
 8 exclamation marks increases slightly over time in two of the newspapers, but
 9 in the others it decreases much more strongly (*Trouw* and *De Telegraaf*). In
 10 background news and hard news, the incidence of exclamation marks decreases,
 11 as well. The incidence of question marks increases in three newspapers, but
 12 decreases in two others (again, *Trouw* and *De Telegraaf*). Finally, the use of
 13 modal adverbs decreases in one of the subgenres, opinion news.

14 The conclusion is, then, that the expected informalization patterns does
 15 occur in direct quotations.

18 4.4 Summary of the results

19 In this paper, the developments of the proportions of subjectivity indicators in
 20 the text “overall” have been compared to the developments of subjectivity indi-
 21 cators in two subsets of the data, namely reporter’s text and direct quotations
 22 of news sources. This comparison shows that separating the different voices in
 23 an analysis of subjectivity is, in fact, crucial. Table 6 gives a summary of the
 24 developments over time of the eleven indicators in the three text types (“over-
 25 all”, reporter’s text, quoted text), disregarding differences between subgenres
 26 and newspapers.

27 With respect to the informalization hypothesis, Table 6 shows that there is a
 28 clear difference between the developments over time in the voice of the reporter
 29 on the one hand, and the developments in the voice of quoted news sources on
 30 the other: the subjectivity of the latter, as expressed in direct quotations, has
 31 increased across all indicators, with the only exception of exclamation marks.
 32 The subjectivity of the reporter, however, has increased in only five of the eleven
 33 indicators (second person singular pronouns, cognitive verbs, question marks,
 34 intensifiers, and modal particles). Moreover, it has decreased in four others (first
 35 person plural pronouns, exclamation marks, modal adverbs, and modal verbs).
 36 In one indicator, there is no change at all (first person singular pronouns). The
 37 results for the deictic elements are ambiguous, because the newspapers and/or
 38 subgenres do not concur in the direction of the change.

1 **Table 6:** Summary of the developments of the 11 indicators in the text overall, in the reporter's
 2 text, and in direct quotations

3	Indicator	Overall	In reporter's 4 text	In direct 5 quotations
6	1. First person singular pronouns	↑	↔	↑
7	2. Second person singular pronouns	↑	↑	↑
8	3. First person plural pronouns	~	↓	↑
9	4. Deictic elements	~	~	↑
10	5. Cognitive verbs	↑	~ ↑	↑
11	6. Exclamation marks	↔	↓	~
12	7. Question marks	~ ↑	~ ↑	~ ↑
13	8. Intensifiers	↑	↑	↑
14	9. Modal adverbs	~	↓	↑
15	10. Modal particles	↑	↑	↑
16	11. Modal verbs	~	↓	↑

17 ↑ = increase, ↓ = decrease, ↔ = no effect of year, so no change, ~ = ambiguous result,
 18 ~ ↑ = ambiguous result but tendency to increase.

20 5 Making statistics concrete: The case of modal 21 adverbs

24 The following trend can be seen in our data: when looking at entire texts, there
 25 is no clear indication of informalization; some features show an increase over
 26 time, whereas others stay the same. In the reporter's text, the trend is even
 27 more mixed, with some features increasing and other features decreasing. In
 28 the direct quotations, we see an overall increase of the subjectivity features.

29 A case in point are the modal adverbs. In the texts "overall," the results of
 30 the analyses are mixed: some newspapers and subgenres show an increase,
 31 some a decrease. In the reporter's text, we see a decrease in most newspapers.
 32 In the direct quotations, the use of modal adverbs shows an increase in two out
 33 of three subgenres. Looking at a particular modal adverb (*natuurlijk*, meaning
 34 'of course'), we see the same pattern: a small increase overall, but a very strong
 35 increase in the direct quotations.

36 This means that examples like (1), where *natuurlijk* occurs in the reporter's
 37 text, are used slightly more often in 1950/1 than in 2002, whereas examples like
 38 (2) (*natuurlijk* in quoted text) are much more common in 2002 than in 1950/1.

1 **Table 7:** Relative frequencies of *natuurlijk* ('of course')
 2 in three subcorpora, by year (per 10,000 words)

3 Subcorpus	1950/1	2002
4 Overall text	2.54	3.01
5 Reporter's text	2.49	1.89
6 Direct quotations	3.09	8.42

7

8

9

10

11

12

13

14

15

16

17

18

19

20

21

22

23

24

25

26

27

28

29

30

31

32

33

34

35

36

37

38

39

40

(1) De mensen voor de etalage zijn het er allen gloeiend over eens: Appel is een oplichter of hij is gek. Het is natuurlijk erg gemakkelijk dat te zeggen. Maar Karel Appel staat niet alleen. In Amsterdam is er een groep schilders, die zo werkt.

The people in front of the shop-window all agree wholeheartedly: Appel is either an impostor, or he is crazy. It is, of course, very easy to say that. But Karel Appel is not alone. In Amsterdam there is a group of painters, that works that way.

(*Algemeen Dagblad*, September 29 1951, section: culture)

(2) Toch zal het wennen worden voor Maxima. “Die Argentijnse mannen zijn natuurlijk vol passie en hartstocht, een laaiend vuur”, denken ze. “Alex is daarentegen een nuchtere man.”

Still it will take some getting used to for Maxima. “Those Argentinian men are, of course, full of passion and fervour, a glowing fire”, they think. “Alex on the other hand is a level-headed man.”

(*Algemeen Dagblad*, February 4 2002, section: domestic news)

30 6 Conclusion

31 Informalization in the text “overall”

32 The results show that analyzing the expression of subjectivity in journalistic dis-
 33 course without accounting for the distinction between reporter's subjectivity and
 34 source subjectivity can lead to misleading results. Indeed, from the analysis of
 35 the indicators in the text “overall”, the conclusion would be that there are
 36 some indicators that provide evidence in favour of informalization, whereas for
 37 other indicators only some newspapers and subgenres show informalization and
 38 others do not. Table 6 shows that for the “overall” text, six indicators demon-
 39 strate an increase (first and second person singular pronouns, cognitive verbs,
 40

1 question marks, intensifiers, and modal particles). These are all indicators for
 2 which the results show an increase in both the reporter's text and the direct quotations.
 3 However, for four other indicators in the text "overall," the results are
 4 inconclusive: there is an effect of the year of publication, but the results for
 5 the five newspapers and three subgenres are so divergent that they cannot be
 6 summarized as a development in either direction, or they cancel each other out
 7 statistically (first person plural pronouns, deictic elements, modal adverbs, and
 8 modal verbs). Table 6 demonstrates that the findings for three of these inconclu-
 9 sive indicators are in fact based on an increase in the direct quotations and a
 10 decrease in the reporter's words. This means that the conclusion based on the
 11 results for the indicators "overall" fails to recognize what is truly happening:
 12 there are two distinct developments; one in the direct quotations, which already
 13 were more subjective than the words of the reporter in 1950/1 and have contin-
 14 ued to become even more subjective; and a different development in reporter's
 15 words, which have informalized only to a degree. These separate developments
 16 in the direct quotations and the reporter's words are obscured in the overall
 17 analysis, because they cancel each other out.

18

19

20 **Informalization in reporter's text and direct quotations**

21

22 Our results show that the informalization hypothesis does not hold equally for
 23 the two subsets of our data. The hypothesis states that in general, lexicogram-
 24 matical indicators of subjectivity show an increase over time.

25

26 In the direct quotations, all these expected increases are observed, although
 27 for the question marks and the modal adverbs, this rise in frequency does not
 28 take place in all newspapers (for question marks) and subgenres (for modal
 29 adverbs). In accordance with the previous findings, exclamation marks have
 30 not increased over time and have even decreased in the newspapers that used
 31 them most in 1950/1 (*De Telegraaf* and *Trouw*). This decrease in the use of excla-
 32 mation marks might be caused by increased prescriptivism, that may have come
 33 up with the professionalization of journalism and the start of professional train-
 34 ing for journalists between 1950 and 2002 (see Vis 2011, Chapter 1, for a discus-
 35 sion of these historical developments in journalism). The use of exclamation
 36 marks in journalistic prose may be deemed too informal to be acceptable in
 37 written text now, whereas this was not the case in 1950/1. Indeed, the present-
 38 day style guide for *de Volkskrant* prescribes to use exclamation marks sparingly
 39 (Van Gessel et al. 1992). Additionally, it states that exclamation marks may be
 40 used for exclamations in quotations, but that as a rule the text should speak
 for itself.

1 The modal verbs, which were not expected to show a rise in frequency in
2 light of the previous findings, did, however, predominantly increase in frequency
3 in the direct quotations, although in hard news in newspapers, this change is
4 not so clear. The difference between this finding and the findings of previous
5 studies with respect to the diachronic change in use of modal verbs might be
6 related to the difference between English and Dutch in the options of expressing
7 modality: in English, there are two candidates for expressing modality in verbs,
8 core modal verbs and semi-modal verbs, whereas in Dutch only the first option
9 is available. As discussed in section 2, studies on changes in English found a
10 decrease in modal verbs and an increase in semi-modal verbs. As the latter
11 option is not available in Dutch, this might (partly) explain the increase in the
12 use of core modal verbs. Another explanation might be found in the fact that
13 the previous diachronic studies, on which our expectations were based, did not
14 systematically account for the distinction between the reporter's words and the
15 direct quotations.

16 With respect to the reporter's words, the expected increases were only
17 observed in second person singular pronouns, cognitive verbs, question marks,
18 intensifiers, and modal particles. First person pronouns (both singular and plural),
19 deictic elements, and modal adverbs did not meet the expectation. In light of
20 the previous findings, exclamation marks and modal verbs were not expected
21 to show an increase, and, in fact, decreased over time. In sum, some indicators
22 in the reporter's words provided evidence in favour of the informalization
23 hypothesis, whereas others display a significant decrease, suggesting that the
24 reporter's words have informalized only to a degree. The increases in second
25 person singular pronouns and question marks suggest that newspapers have
26 become more inclined to express interactive indicators over time. In contrast,
27 the decreasing occurrence of deictic elements and first person plural pronouns,
28 and the stable frequency of first person singular pronouns might mean that
29 newspapers have become more reluctant over time to explicitly refer to the
30 here-and-now and to themselves as institutions or to the community or nation
31 that they are part of. Similar to the results for the direct quotations, the differences
32 between the results of this study for the reporter's text and the findings of
33 previous studies might be caused by the fact that the earlier findings were based
34 on analyses that did not systematically take the distinction between reporter's
35 text and direct quotations into account.

37 7 Discussion

38
39
40 To what extent is informalization dependent on subgenre and on different news-
papers? It is noteworthy that the newspapers and subgenres are involved in a

1 two- or three-way interaction with the year of publication in a markedly high
 2 number of cases. Newspaper is involved in an interaction with year of publica-
 3 tion in 25 out of 33 times (11 indicators times 3 for text “overall”, reporter’s text,
 4 and direct quotations) and subgenre in 22 out of 33 cases. This means that for
 5 the majority of indicators, there are significant differences between the news-
 6 papers and/or between the subgenres. This is not necessarily a difference with
 7 respect to the direction of the development; in fact, it is striking how often the
 8 newspapers and subgenres concur in the direction of the change, both in the
 9 reporter’s text and the direct quotations. Rather, they vary in the extent to which
 10 they change over time.

11 In some cases, the newspapers and subgenres are not involved in an inter-
 12 action with year; for example, in first person singular pronouns in the text
 13 “overall” and in the reporter’s words, there are no significant differences
 14 between newspapers or between subgenres with respect to their development
 15 over time. However, even if there is no interaction with year, the newspapers
 16 and subgenres usually still have an effect, either in a two-way interaction with
 17 each other or separately as a main effect. This means that in those cases, there
 18 is a difference between newspapers and/or subgenres, but this difference does
 19 not change over time.

20 For only four indicators, the newspapers and/or subgenres show no differ-
 21 ence at all. Interestingly, for each of these four indicators, this concerns the
 22 incidences in direct quotations: there is no effect of newspaper on deictic ele-
 23 ments and cognitive verbs in direct quotations, and no effect of subgenre on
 24 deictic elements, question marks, and intensifiers in direct quotations.

25 Although this shows that there are significant differences between news-
 26 papers and subgenres in their behaviour with respect to the use of subjective
 27 elements, it is difficult to make general statements about the character of these
 28 differences, since there are so many different indicators and interactions in-
 29 volved in the analyses. There is no newspaper that shows a clear consistent
 30 development across all indicators and types of voice (reporter or direct quota-
 31 tion). Note that there is also no clear difference between tabloids (*De Telegraaf*,
 32 *Algemeen Dagblad*) versus quality papers (*NRC Handelsblad*, *de Volkskrant*,
 33 *Trouw*). As mentioned before, trends toward tabloidization are less strong in
 34 the Netherlands compared to, for example, the United Kingdom. Similarly,
 35 although subgenre was involved in various interactions with year of publication,
 36 the pattern was too irregular to reach firm conclusions about the role of sub-
 37 genre in the informalization process.

38 In conclusion, this study provides evidence that it is crucial to distinguish
 39 between reporter’s subjectivity and subjectivity of news sources quoted in the
 40 text. With respect to the informalization hypothesis, it seems that the reporter’s

1 text has informalized only to a degree: some indicators show evidence in favour
 2 of the informalization hypothesis, whereas others appear to be evidence against
 3 it. As suggested above, the results for the reporter's text might mean that news-
 4 papers tend to express more interactive indicators over time, whereas they
 5 became more reluctant to explicitly refer to themselves and to the here-and-now.
 6 Direct quotations, however, show strong evidence of informalization. This devel-
 7 opment seems to be relatively independent of the subgenre that is investigated.

8 This study contributes to our understanding of developments in journalistic
 9 language use. News reports have often been treated as a single register. We have
 10 shown how such a limitation can be overcome using corpus analysis. By using a
 11 sophisticated type of statistics, we were able to paint a detailed picture of a com-
 12 plex diachronic development using a large set of linguistic features. We also
 13 believe to have added to the existing body of knowledge which mainly focuses
 14 on English language use, by describing patterns in Dutch journalistic language
 15 use, occurring in the major national newspapers of the Netherlands.

18 References

- 20 Bekker, Birgit. 2006. *De feiten verdraaid. Over tekstvolgorde, talige markering en sprekerbe-*
 21 *trokkenheid* [Twisted facts. On discourse order, linguistic marking and speaker involve-
 22 ment]. PhD dissertation/ Tilburg: University of Tilburg.
- 23 Biber, Douglas. 1988. *Variation across speech and writing*. Cambridge: Cambridge University
 24 Press.
- 25 Biber, Douglas. 2004. Historical patterns for the grammatical marking of stance: A cross-
 26 register comparison. *Journal of Historical Pragmatics* 5(1), 107–136.
- 27 Biber, Douglas & Edward Finegan. 1989. Drift and the evolution of English style: a history of
 28 three genres. *Language* 65, 487–517.
- 29 Biber, Douglas & Edward Finegan. 2001. Diachronic relations among speech-based and written
 30 registers in English. In Susan Conrad & Douglas Biber (eds.), *Variation in English: Multi-*
 31 *Dimensional studies*. 66–83. London: Longman.
- 32 Biber, Douglas & Edward Finegan, Dwight Atkinson, Ann Beck, Dennis Burges, & Jená Burges.
 33 1994. The design and analysis of the ARCHER Corpus: A progress report. In Merja Kytö,
 34 Matti Rissanen & Susan Wright (eds.), *Corpora across the centuries (Proceedings of the*
 35 *First International Colloquium on English Diachronic Corpora)*. 3–6. Amsterdam: Rodopi.
- 36 Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad, & Edward Finegan. 1999. *The*
 37 *Longman Grammar of Spoken and Written English*. London: Longman.
- 38 Bolinger, Dwight. 1972. *Degree words*. The Hague: Mouton.
- 39 Deuze, Mark. 2005. Popular journalism and professional ideology: Tabloid reporters and editors
 40 speak out. *Media, Culture & Society* 27(6), 861–882.
- Fairclough, Norman. 1994. Conversationalization of public discourse and the authority of the
 consumer. In Russel Keat, Nigel Whiteley & Nicholas Abercrombie (eds.), *The Authority of*
the Consumer. 253–268. London: Routledge.

- 1 Fairclough, Norman & Ruth Wodak. 1997. Critical Discourse Analysis. In Teun A. van Dijk (ed.),
2 *Discourse Studies: A Multidisciplinary Introduction. Volume 2: Discourse as Social Interac-*
3 *tion.* 258–284. London: Sage.
- 4 Haeseryn, Walter, Kirsten Romijn, Guido Geerts, Jaap de Rooij, & Maarten van den Toorn. 1997.
5 *ANS. Algemene Nederlandse Spraakkunst* [General Dutch grammar], 2nd edn. Groningen:
6 Martinus Nijhoff.
- 7 Hundt, Marianne & Christian Mair. 1999. “Agile” and “uptight” genres: the corpus-based
8 approach to language change in progress. *International Journal of Corpus Linguistics* 4(2).
9 221–242.
- 10 Knobloch, Sylvia, Grit Patzig, & Anne-Maria Mende 2004. Affective news: Effects of discourse
11 structure in narratives on suspense, curiosity, and enjoyment while reading news and
12 novels. *Communication Research* 31(3). 259–287.
- 13 Langacker, Ronald. 1990. Subjectification. *Cognitive Linguistics* 1. 5–38.
- 14 Leech, Geoffrey, Marianne Hundt, Christian Mair, & Nicholas Smith. 2009. *Change in contem-*
15 *porary English.* Cambridge: Cambridge University Press.
- 16 Lyons, John. 1982. Deixis and subjectivity: Loquor, ergo sum. In R.J. Jarvella & Wolfgang Klein
17 (eds.), *Speech, Place, and Action: Studies in Deixis and Related Topics.* 101–125. New York:
18 Wiley.
- 19 Mair, Christian. 2006. *Twentieth-century English.* Cambridge: Cambridge University Press.
- 20 Pasma, Trijntje. 2001. *Metaphor and register variation: The personalization of Dutch news*
21 *discourse.* PhD Dissertation. Amsterdam: VU University.
- 22 Pearce, Michael. 2005. Informalization in UK party election broadcasts: 1966–97. *Language &*
23 *Literature* 14(1). 65–90.
- 24 Precht, Kristen. 2000. *Patterns of stance in English.* Unpublished dissertation, Northern Arizona
25 University, Arizona, USA.
- 26 Rafiee, Afrooz. In preparation. The discourse structure of Iranian child abuse news.
- 27 Scheibman, Joanne. 2002. *Point of view and grammar: Structural patterns of subjectivity in*
28 *American English conversation.* Amsterdam: John Benjamins.
- 29 Steen, Gerard. 2003. Conversationalization in discourse: Stylistic changes in editorials of The
30 Times between 1950 and 2000. In Luuk Lagerwerf, Wilbert Spooren & Liesbeth Degand
31 (eds.), *Determination of Information and Tenor in Texts: Multidisciplinary Approaches to*
32 *Discourse.* 115–124. Münster: Nodus Publikationen.
- 33 Steen, Gerard. 2011. Genre between the humanities and the sciences. In Marcus Callies, Wolfram
34 R. Keller, and Astrid Lohöfer (eds.), *Bi-Directionality in the Cognitive Sciences: Avenues,*
35 *challenges, and limitations.* 21–41. Amsterdam etc.: Benjamins.
- 36 Thompson, Geoffrey. & Susan Huston. 2000. Evaluation: an introduction. In Susan Huston &
37 Geoffrey Thompson (eds.), *Evaluation in text: Authorial stance and the construction of*
38 *discourse.* 1–27. Oxford: Oxford University Press.
- 39 Traugott, Elisabeth. 1989. On the rise of epistemic meanings in English: an example of subjec-
40 tification in semantic change. *Language* 57. 33–65.
- van den Bosch, Antal, Bertjan Busser, Walter Daelemans, & Sander Canisius. 2007. An efficient
memory-based morphosyntactic tagger and parser for Dutch. In Frank van Eynde, Peter
Dirix, Ineke Schuurman & Vincent Vandeghinste (eds.), *Selected Papers of the 17th*
Computational Linguistics in the Netherlands Meeting. 99–114. Leuven.
- van Gessel, Han, Jan Kees Hulsbosch, Henk Huurdeman, Bas van Kleef, Kees Los, & Bert
Vuijsje. (eds.) 1992. *De Volkskrant Stijlboek* [De volkskrant style guide]. 's-Gravenhage:
Sdu Uitgeverij.

- 1 van Vree, Frank, Chris Vos, Huub Wijffes & Jo Bardoel. 2003. Het ontstaan van een politiek-
2 publicitair complex 1960–2002 [The rise of a political and press-related complex 1960–
3 2002]. In RMO, *Medialogica. Over het krachtenveld tussen burgers, media en politiek*
4 [Media logic. On the power relations between citizens, media and politics]. 67–98. Den
5 Haag: Raad voor Maatschappelijke Ontwikkeling.
- 6 Vis, Kirsten. 2011. *Subjectivity in news discourse: A corpus linguistic analysis of informaliza-*
7 *tion*. PhD Thesis. Amsterdam: VU University Amsterdam. Available from: [http://dare.uvu.](http://dare.uvu.vu.nl/handle/1871/19775)
8 [vu.nl/handle/1871/19775](http://dare.uvu.vu.nl/handle/1871/19775)
- 9 Vis, Kirsten, Wilbert Spooren & José Sanders. 2010. Using RST to analyze subjectivity in text
10 and talk. In Elzbieta Tabakowska, Michal Choinski & Lukasz Wiraszka (eds.), *Cognitive*
11 *Linguistics in Action: From Theory to Application and Back*. 293–316. Berlin: Mouton de
12 Gruyter.
- 13 Vis, Kirsten, José Sanders & Wilbert Spooren. 2012. Diachronic changes in subjectivity and
14 stance – A corpus linguistic study of Dutch news texts. *Discourse, Context & Media* 1 (2–3).
15 95–102.
- 16 Westin, Ingrid & Christer Geisler. 2002. A multi-dimensional study of diachronic variation in
17 British newspaper editorials. *ICAME Journal* 26. 133–152.
- 18 Wiebe, Janyce. 1994. Tracking point of view in narrative. *Computational Linguistics* 20. 233–
19 287.

20 Appendix: Subjectivity lexicon

21 For the analysis, a lexicon of subjectivity elements was created, based on insights
22 from the literature.

- 23 1. First and second person pronouns [a.o. Scheibman 2002]
24 All lexical items tagged as first or second person pronoun, either personal or
25 possessive, were coded as subjective. Additional codes were set to distin-
26 guish between person (first or second person) and number (singular or
27 plural).
28
- 29 2. Modal adverbs [a.o. Biber, Johansson, Leech, Conrad, and Finegan 1999]
30 33 frequent modal adverbs were counted per text. Some of these items can
31 only be used as adverbs, e.g. *allicht* ('of course'), *helaas* ('unfortunately')
32 and *misschien* ('maybe'). However, some of the items are in fact adjectives
33 that can be used either adjectively or adverbially, such as *natuurlijk*
34 ('natural'/'of course') and *mogelijk* ('possible'/'possibly'). The annotation
35 scheme only identified uninflected forms of these items as modal adverb.
36 The instances matching the identification patterns were tagged as subjective
37 category 'modal' and subcategory 'adverb'.
- 38 3. Modal particles [a.o. Precht 2000]
39 Modal particles are uninflected words that reflect the attitude of the speaker.
40 Many of the words functioning as a modal particles have other functions as

1 well. In the set used in this study, only those modal particles were included
 2 that could automatically be recognized as such on the basis of its part of
 3 speech tags in combination with (the part of speech tags of) the immediate
 4 context. For example, *maar* ('but') as an adverb was annotated as modal
 5 particle, whereas instances of *maar* as a conjunctive were not; the adverb
 6 *al* ('already') was only annotated as a modal particle when it was not
 7 followed by an adjective, adverb, numeral/quantifier or pronoun; and the
 8 lexical item *pas* ('only') tagged as an adverb was annotated as a modal par-
 9 ticle except when it was preceded by *van* ('of') or *te* ('to'), which happened
 10 rather frequently in the 1950s corpus.

11 4. Intensifiers [Haeseryn et al. 1997]

12 The lexicon contains three sets of intensifiers, grouped with respect to the
 13 conditions necessary for the identification of the intensifier. The items in
 14 the first set were words like *bijna* ('almost'), *echt* ('really'), and *nogal*
 15 ('rather'). The second set consists of prefixes that all intensify the sense of
 16 the words they are added to. For instance, *sterk* means 'strong', but *oersterk*
 17 means 'exceedingly strong'. Finally, the third set consists of adverbs that are
 18 only used as intensifiers when the following (or preceding, in the last two
 19 cases) item is an adverb, adjective, or pronoun; it contains elements like
 20 *amper* ('hardly'), *redelijk* ('reasonably'), and *verschrikkelijk* ('terribly').

21 5. Modal verbs [Haeseryn et al. 1997]

22 The subjectivity lexicon adopted the list of modal verbs presented by Haeseryn
 23 et al. (1997):

24 *behoeven, blijken, dunken, heten, hoeven, lijken, kunnen, moeten, mogen,*
 25 *schijnen, toeschijnen, voorkomen, willen, zullen*

26 All forms of these verbs (both finite and other verb forms) were annotated as
 27 modal verbs.

28 6. Cognitive verbs [Biber 2004]

29 A list of frequently used cognitive verbs was compiled. Following Biber
 30 (2004: 111), a liberal approach was adopted in defining the set of cognitive
 31 verbs, including "any form that can be interpreted as expressing author/
 32 speaker attitude and assessment of likelihood". The resulting list contains
 33 the following 52 verbs, including non-factive verbs commenting on how
 34 information is communicated, e.g. *antwoorden* ('answer'), attitudinal verbs,
 35 e.g. *hopen* ('hope'), factive verbs, e.g. *weten* ('know') and likelihood verbs,
 36 e.g. *betwifelen* ('doubt').

37 The annotation scripts identified the cognitive verbs on the basis of the
 38 lemma information in the corpus. This ensured that all different forms of
 39 the verbs were annotated as subjective. Additionally, identifying the cogni-
 40 tive verbs by their lemma has the added benefit that versions of the verbs

1 that are now considered archaic, such as *zeide* ('said'), are annotated as
2 cognitive verbs.

3 7. Exclamations

4 Both exclamations and imperatives are usually marked by the same punctua-
5 tion mark, the exclamation mark. The scheme for subjectivity annotation used
6 this symbol to identify both features and added a subjectivity code to the XML
7 tags for the exclamation marks.

8 8. Questions

9 The identifying patterns identified questions by the punctuation mark that
10 typically marks them, the question mark. The annotation added a subjectivity
11 code to each question mark in the corpus.

12 9. Deictic elements [Bekker 2006]

13 The lexicon contains a limited set of deictic elements, such as *hier* ('here'),
14 *nu* ('now'), *gisteren* ('yesterday'), *vandaag* ('today'), *morgen* ('today'). These
15 elements were identified through their lemma information in the corpus
16 and were annotated with a subjectivity code.

17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40