Addressing the student: Voice elements in educational texts

Nina L. Sangers[1], Jacqueline Evers-Vermeul[1] and Hans Hoeken[1]

[1] *Utrecht Institute of Linguistics OTS, Utrecht University*

Abstract Voice elements are those elements of educational texts that authors use to interact with students, such as questions, evaluations, or direct address forms ('you'). These elements are intended to enhance students' engagement and comprehension, but we know little about the extent to which they are used in present-day educational texts. Using a corpus of Dutch biology, geography, and history texts for grade 5 and grade 8 ($N = 1055$), this study shows that voice elements are barely differentiated over grade levels. Conversely, voice elements are generally varied over school subjects, as they are less frequent in history texts – which convey readily imaginable and relatable content – compared to biology and geography texts – which discuss less relatable content for which students need to exert more effort to connect it to their own world. This finding suggests that authors of educational texts have intuitions about the conditions under which voice elements are a desirable attribute.

Keywords voice elements, educational texts, author-student interaction, distance, relatability educational content, quantitative corpus-based analysis

1 Introduction

Despite rapid technological developments, textbooks are still the predominant instructional medium in present-day Dutch classrooms (Woldhuis et al., 2018). Since most learning is accomplished through reading of educational texts, students' success in school is highly dependent on their understanding of these texts. However, many Dutch students consider their educational texts too boring to read and/or too difficult to

understand (Dood et al., 2020; Gubbels et al., 2017; Gubbels et al., 2019; Inspectorate of Education, 2017, 2020, 2021). Making educational texts more engaging has been argued to be an adequate means to solve these readability issues: students who are engaged in the active processing of educational texts are likely to be more attentive to the to-be-learned information presented in these texts, which stimulates deeper understanding and better learning of this information (cf. Beck et al., 1995; Brozo et al., 2007; Guthrie, & Wigfield, 2000; Hidi, 2001; Sadoski, 2001; Schraw & Lehman, 2001).

Beck et al. (1995) were the first to introduce *voice* as a strategy to make educational texts more engaging. They developed a notion of voice that consists of three themes: activity, orality, and connectivity. Activity involves making an educational text more dynamic by using verbs that represent concrete action, and by describing the immediacy of events and characters' responses to these events. Orality covers the use of conversational language, including colloquial expressions, emphatics, and explicit dialogue. Connectivity involves the emphasizing of relationships, such as addressing the reader directly, drawing connections between events and characters' responses, and vitalizing interrelationships among characters within the text.

Based on these themes, Beck et al. (1995) created voiced and non-voiced versions of a history text, for which a coherent and a non-coherent version were already developed in previous empirical research (cf. Beck et al., 1991; McKeown et al., 1992).[1] In an effect study, Beck et al. (1995) established that fourth graders achieved better comprehension for the text version that both contained voice and was coherent than for versions of this text that lacked either one or both of these aspects. These findings led Beck et al. (1995) to tentatively conclude that voice enhances students' comprehension of educational texts in coherent contexts.

Although we acknowledge the potential of voice elements in making educational texts more engaging, and consequently better comprehensible, we believe that Beck et al.'s (1995) notion of voice is too wide-ranging, and hence, is in need of specification. We propose to define voice solely in terms of textual elements that directly relate to the author, in the current paper of educational texts. While the activity of writing educational texts is in itself monologic, its communicative structure is dialogic, being the product of a reciprocal relationship between author and students (cf. Nystrand, 1986; Vološinov, 1973).[2] The author can make this relationship explicit by using specific textual elements such as speech acts (see Section 2) to openly interact with students. For instance, by asking students a question, the author of the history text in (1) invites students to actively think about the educational content and helps them to relate this content to their personal lives, thereby decreasing the distance between them and this content (cf. Sangers, 2022; Sangers et al., 2020).[3]

(1)   *Do you prefer your clothes to look nice over them being comfortable?* People used to consider their clothes mainly important for protecting their bodies against the

weather or against scratches of branches. This means that clothes had to be robust and warm.

(*Wijzer! Geschiedenis*, history grade 5, p. 16)

Author-initiated dialogical elements, such as questions, form an important subset of Beck et al.'s (1995) catch-all interpretation of the concepts of connectivity and orality. For connectivity, however, in describing the relationship between text and reader, Beck et al. (1995) do not distinguish between author-initiated elements, which generate a layer on top of the text's content (cf. Smolkin et al., 2008), and elements that more directly relate to this content, such as characters' responses to events and their interrelationships. Similarly, for orality, Beck et al. (1995) do not make a distinction between the conversational language expressed by the text's author to initiate an interaction with students, and conversations that are text internal, among characters. These latter kinds of elements do not directly relate to the educational text's author, but rather have to do with narrativity (cf. Toolan, 2001; Sangers, 2022; Sangers et al., 2020, 2021). We believe that author-initiated dialogical elements and narrative elements are two separate strategies to enhance students' engagement and text comprehension, and hence, should be defined and explored independently (Sangers, 2022; Sangers et al., 2020).

In this study, we focus on the dialogical elements of educational texts that are initiated by these texts' author, which we define as the *voice*. While the inclusion of voice elements in educational texts seems to be an attractive strategy to enhance students' engagement and comprehension, we know little about the extent to which these elements are used in present-day educational texts. To gain more insight into the distribution of voice elements in educational texts, we conducted a quantitative corpus-based analysis, focusing on current Dutch practice. Our analysis was guided by the following research question:

How and when are voice elements currently being used in Dutch educational texts?

Before we turn to the method of our corpus-based analysis (Section 5), we describe the different ways in which voice elements can be expressed in Dutch educational texts (Section 2), and formulate hypotheses concerning potential variation in educational publishers' distribution of these elements over school subjects (Section 3) and/or grade levels (Section 4).[4]

2 The role of voice elements in Dutch educational texts

Voice elements surface in various ways in Dutch educational texts (cf. Sangers, 2022; Sangers et al., 2020). Besides asking questions, as in (1), the author of an educational text can use imperatives to encourage students to do something, as in (2), or instruct them to imagine a situation sketched in the text, as in (3).

(2)   *Put on* a helmet and also try to find a pair of goggles. *Hang upside down* on the climbing frame for a minute. Occasionally *straighten up* your upper body.
       (*Argus Clou*, history grade 5, p. 56)

(3)   Standing under a powerful jet of water in the shower in the morning is wonderful! *But imagine what would happen if* there were sand in the water: your skin would be sandblasted, and that is less pleasant. Let this exactly be what rivers do to rocks ...
       (*BuiteNLand*, geography grade 8, p. 34)

In addition, the author can highlight important or interesting educational content by adding an exclamation mark at the end of the sentence, as in (4). Such exclamations are intended to direct students' attention to specific parts of the educational text, and to make them consider why this particular information stands out.

(4)   *Some seed plants can reproduce without seeds!* How do they do that? By growing a new plant from a stem, a root, a tuber, or a bulb.
       (*Argus Clou Natuur & Techniek*, biology grade 5, p. 51)

Another way in which the author can highlight educational content is by giving evaluations. Evaluations give students insight into the author's attitudes, feelings, values of judgement, or expectations about the text's content (cf. Conrad & Biber, 2000), and implicitly invite them to assess whether they agree with what is being said. For instance, the evaluation in (5) states that the author considers the European Union very special. It is left to students to decide whether they agree with this evaluation or not.

(5)   You belong to it: the European Union. As many as 28 different countries follow the same rules and laws. They collaborate at an economic and political level. Some countries are rich, others less so. People speak different languages, have different customs and yet things usually go well. *Very special, this European Union!*
       (*Grenzeloos*, geography grade 5, p. 34)

Questions, imperatives, exclamations, and evaluations expressed by the author of an educational text can be defined as speech acts, as they invite or urge students to actively deal with or think about the educational content (cf. Austin, 1962; Searle, 1969, 1975; Houtkoop, & Koole, 2000; Pander Maat, 2002). Such speech acts underline the interactive nature of educational texts and contribute both to the level of orality and to the level of connectivity of these texts (Beck et al., 1995).

   Besides speech acts, the author can use first- and second-person pronouns to start or continue a conversation with students. The author can use the second-person singular pronouns *you* and/or *your*(*self*) to directly address students as individuals, as in (5) and

(6), and/or the first-person plural pronouns *we* and/or *our*(*self*) to address students as a group to which the author belongs as well, as in the final sentence of (6).

(6) From space *you* have to look carefully to see the small Netherlands. But if *you* know where the North Sea is located, *you* can find it quickly. The North Sea determines *our* climate. *We* have a moderate maritime climate here.
(*Argus Clou*, geography grade 5, p. 21)

The personalizing function of personal pronouns can result in a higher level of student involvement and/or identification (cf. De Hoop, & Tarenskeen, 2012; Wechsler, 2010), as it enables students to relate the educational content to their own lives. Several researchers (cf. Deringer et al., 2015; Gast et al., 2015) claim that this effect holds, irrespective of whether second person pronouns are applied in a personal way (referring exclusively to the addressee) or a more generic, impersonal way (referring to a more comprehensive set of referents; compare *You know this climate all too well: you live in it!* to *You can find this climate for instance in the Netherlands*). In their view, impersonal, generalizing instances of *you* are canonical instances of *you* occurring in sentences expressing a generalization about a category to which the addressee happens to belong. When addressees are a member of the category in question, empathy is intrinsic; when they are not, empathy is created through simulation (Gast et al., 2015, p. 161). In line with this view, Orvell et al. (2020) have shown that ideas expressed with impersonal, generic *you* result in increased resonance, a sense of connection, compared to ideas expressed without such a pronoun. All in all, educational authors have a range of voice elements at their disposal.

3 The role of voice elements for different school subjects

It is plausible that educational publishers choose to differentiate their use of voice elements over school subjects. As voice elements are supposed to decrease the distance between students and the educational content, they seem to be most beneficial for school subjects whose educational content is abstract and less personally relatable, such as biology and geography (Sangers, 2022; Sangers et al., 2021). In human-oriented school subjects, such as history, people regularly feature in the educational content, alongside their goals, experiences, and actions (Sangers, 2022; Sangers et al., 2021). These individuals enable identification, and their presence can help students to view the educational content from a different perspective, making the to-be-learned information easily imaginable and relatable (cf. Hidi, 2001; Kuijpers, 2014). In this respect, historical figures act as go-betweens: they help to decrease the distance between the educational content and students. If such mediators are absent, as is the case in texts for less human-oriented school subjects, students need to exert more effort to connect the to-be-learned information to their own world (cf. Sangers, 2022; Sangers et al., 2021). In these cases, the author

can extend a helping hand by acting as an alternative mediator in bridging the distance between the educational content and students (cf. Nolen, 1995). Hence, we hypothesize that publishers use voice elements more frequently in biology and geography texts than in history texts.

At this point, however, it is important to take into account that the school subject geography not only includes topics that are not human-oriented. Besides texts that focus on physical geography topics, such as (7), geography textbooks include content that more closely relates to humans, such as (8).

(7)   Because the earth's axis is tilted, the sun in the Northern Hemisphere rises high above the horizon in the summer. In the far north, it does not even set at night. From May 20 to July 22, it is light here all day and night. We call this the midnight sun. In the winter, the sun does not rise at all in this area. This is called the polar night.
(*Argus Clou*, geography grade 5, p. 19)

(8)   Like the Netherlands, the European Union has its own government: the European Commission (EC). The European Commission is based in Brussels. The EC consists of 27 commissioners, a kind of cabinet minister. They are appointed by the government of their own country. Each member state provides one commissioner, irrespective of how big or small the country is.
(*Argus Clou*, geography grade 5, p. 88)

Following the line of reasoning above, we hypothesize that publishers use voice elements more frequently in physical geography texts than in human geography texts, because the educational content these texts aim to convey is more abstract and less personally relatable. In addition, we hypothesize that voice elements are less frequently used in human geography texts than in history texts, because these texts tend to focus on general tendencies rather than specific events, and on groups rather than individuals. As such, we expect human geography texts to occupy an intermediate position between biology and physical geography texts on the one hand, and history texts on the other hand. Taken together, we hypothesize that the distribution of voice elements in Dutch educational texts depends on school subject in the following way:

Hypothesis 1: history < human geography < physical geography = biology

4 The role of voice elements for different grade levels

It is also plausible that educational publishers use voice elements differently depending on grade levels. On the one hand, it has been widely agreed that educational texts should

become more complex as students' reading proficiency develops (Snow, 2002). Ideally, this reflects a progression in the course of a school career from relatively simple to more sophisticated and challenging texts (Brabham & Villaume, 2002; Shanahan et al., 2012), enabling students to progressively improve their functional literacy, and teaching them how to read texts that gradually become more abstract and less related to their personal experiences (Schleppegrell, 2004). This involves a development from texts with a personal writing style to texts that exhibit a more distant way of writing (Committee Meijerink, 2009; Evers-Vermeul & Holtermann, 2013). Since voice elements facilitate a direct, 'here and now' interaction between the author of the educational text and students, they contribute to a personal way of writing, and can therefore be used to differentiate between simple and more complex educational texts. Hence, publishers may gradually decrease their use of voice elements in educational texts, with more of these elements present in texts for primary education than in those for secondary education. Taking potential differences between school subjects into account, we expect this decrease to mainly occur in texts for the less relatable school subjects biology and physical geography. That is, it is plausible that a decrease in voice elements is less prominent or even absent in the human-related school subjects history and human geography, since the number of voice elements in these subjects might already be limited, leading to potential floor effects.

On the other hand, the educational content in texts for secondary education tends to be more complex than that in primary education texts (Committee Meijerink, 2009; Hidi, 2000). Accordingly, in secondary education texts, voice elements might be helpful in bridging the increasing gap between the to-be-learned information and students (cf. Sangers, 2022; Sangers et al., 2020). Hence, publishers may use voice elements more frequently in texts for secondary education compared to those for primary education. Again considering the potential differences between school subjects, we expect this increase to be more visible in texts for history and human geography than in those for biology and psychical geography, since in the latter subjects, the number of voice elements might reach ceiling effects.

In order to shed light upon this discrepancy in the predicted use of voice elements for different grade levels, our corpus-based analysis focused on texts for grade 5 of primary education and for grade 8. While grade 5 students (age 10–11) have acquired the basic reading skills required for a deep understanding of texts, grade 8 students (age 13–14) need to be able to read more challenging texts, particularly in pre-university education (Committee Meijerink, 2009).[5] Taken together, we formulate two contrasting hypotheses for the influence of grade level on the distribution of voice elements in Dutch educational texts:

> Hypothesis 2A – progression simple-challenging: grade 5 > grade 8
> Hypothesis 2B – bridging student-content gap: grade 5 < grade 8.

5 Method

In this section, we describe the material selection (Section 5.1), method of analysis (Section 5.2), inter-annotator agreement (Section 5.3), and method of statistical analysis (Section 5.4).

5.1 Material selection

5.1.1 *Textbook selection*

We selected materials for grade 5 and for grade 8, in which students are still taking classes in all school subjects under investigation. For grade 8, we focused on pre-university education (*vwo 2* in the Dutch system), because this is the most advanced Dutch secondary education level, which allowed us to maximize the chance of finding differences with materials developed for grade 5.

We selected educational texts from textbooks published by five well-known Dutch educational publishers. For grade 5, one textbook was selected per subject per publisher. Since physical and human geography topics are combined in Dutch geography textbooks, a total of fifteen grade 5 textbooks were selected. For grade 8, only three out of five publishers also distributed textbooks at a pre-university level. All three did so for geography and history, while only two published a biology textbook, leading to eight grade 8 textbooks being selected. See Appendix B for an overview of all twenty-three textbooks.

5.1.2 *Chapter selection*

For history and biology, we selected one chapter per textbook. Per geography textbook, two chapters were selected: one for human geography and one for physical geography. This resulted in the selection of thirty-one chapters. We strived for thematic overlap per subject, both within and between grade levels, in order to counter potential distribution biases caused by topic selection as much as possible. Thematic overlap was established by comparison of keywords. See Appendix C for a more elaborate justification of the selected themes.

5.1.3 *Texts*

From the aforementioned chapters, we selected texts that included educational content and/or background information. Texts that originated from sources other than the educational textbook itself, such as newspaper articles or non-fictitious testimonials from historical figures, were excluded from the corpus. A text was taken to be a unit of at least three sentences that 1) belonged to a marked text box, and/or 2) was grouped under a subheading (blank lines did not mark the beginning of a new text). In those few cases in which these criteria did not suffice, we looked at font characteristics in order to make a final decision. Table 1 shows the number of texts per school subject and level. In total, the corpus consisted of 1055 texts.

Table 1 Number of texts per school subject and level

Subject	Grade 5	Grade 8	Total
Biology	161	87	248
Geography – physical	137	121	258
Geography – human	135	122	257
History	152	140	292
Total	585	470	1055

5.2 Method of analysis

For each text, excluding its heading, we applied binary coding, listing whether each of the following types of voice elements were present or not: 1) questions; 2) imperatives; 3) encouragements to imagine; 4) exclamations; 5) evaluations; 6) instances of the Dutch equivalents of the second-person singular pronoun *you*; 7) instances of the Dutch equivalent of the first-person plural pronoun *we*. To be more precise about the sixth category: as the use of Dutch *je* 'you' is very flexible (Andeweg et al., 2013; Hogeweg, & De Hoop, 2015), we looked for instances of *je* (unaccented/generic 'you'), *jij* (accented/specific subject 'you'), *jou* (accented/specific object 'you'), the possessive *je/jouw* 'your', and/or the reflexive *jezelf* 'yourself'. As *je* can be used in both a personal and an impersonal, generic way (see Section 2), this means that we looked at instances in which students are addressed directly as individuals, or as part of a group. For the seventh category, we included Dutch *we* (unaccented/generic 'we'), *wij* (accented/specific 'we'), *ons/onze* 'our' (possessive), and/or *onszelf* 'ourselves' (reflexive).

A phrase could contain more than one type of voice elements. For instance, in (1), both the question and the pronoun *you* (*Do you prefer your clothes to look nice over them being comfortable?*) were positively scored, while in (5), the evaluation (*Very special, this European Union!*) was also taken to be an exclamation. Voice elements in lines uttered by characters were only scored if these characters directly interacted with the student, and never if they were talking to each other. For instance, in (9), the student is directly addressed with *you(r)*, leading to a positive score.

(9)   Guten Tag! I am Matthias Sammer. I live in Berlin and I enjoy being *your* guide while *you* are exploring my country.
       (*BuiteNLand*, human geography grade 8, p. 17)

For imperatives, we excluded references to sources and other pages of the textbook, because the formulation of these references was arbitrary: some publishers used imperatives in their references (*see source 1*), while others only mentioned the source (*source 1*).

Table 2 Inter-annotator agreement (Cohen's kappa and % agreement) per voice element

Voice element	Cohen's kappa	% agreement
Speech acts		
Questions	.91	97
Imperatives	.97	99
Encouragements to imagine	.60[6]	97
Exclamations	.91	98
Evaluations	.59	89
Pronouns		
You	.98	99
We	.97	99

For evaluations, a 0-bias was followed: only statements that were clearly evaluative received a positive score. In addition, the following instances did not count as evaluations: 1) intensifiers (*it was extremely warm*); 2) substantiated judgements (*the relative location is good: it is not far away and easy to reach*); 3) generally received judgements (*the climate in Spain is good*); 4) fixed expressions (*set a good example*); 5) judgements ascribed to others (*for many companies, the Netherlands was an attractive place of business*); and 6) instances of epistemic stance (how certain the author is about the given information, e.g., *apparently*) or style stance (the manner in which the information is presented, e.g., *briefly*).

5.3 Inter-annotator agreement

For considerations of reliability, 15% of the corpus (N = 166) was coded by a second, independent annotator (cf. Neuendorf, 2002). This sample was randomly compiled for each school subject and grade level. Before the second annotator coded the sample, she engaged in a training phase to make her familiar with the procedure and the elements under investigation. The inter-annotator agreement was moderate to almost perfect (.59 < κ > 1.00) (cf. Landis & Koch, 1977), as shown in Table 2.

    The annotators discussed and resolved disagreements in their analyses to reach a final dataset. As the somewhat lower kappa-score for evaluations indicates, annotators relatively often disagreed about the presence of evaluations, for instance about the subjectivity of the word *normal* in (10). It was reasoned that although this word is somehow evaluative (what is considered normal can vary from person to person), it is not used in an argumentative way here.

(10) The disease [AIDS] cannot be transmitted by sneezing, coughing, kissing or by using each other's cutlery. You can therefore continue to interact with AIDS patients or HIV positive people in a normal way without danger.
(*Biologie voor jou*, biology grade 8, p. 188)

5.4 Statistical analysis

The final dataset was analyzed using R version 3.6.1 (R Core Team, 2019). The analyses were completed via generalized linear mixed models, using the packages haven (Wickham & Miller, 2019), lme4 (Bates et al. 2015), emmeans (Lenth, 2019), and ggplot2 (Wickham, 2016). The fixed factors *Subject* and *Level*, and their interaction, were added to the models in a stepwise manner. Because some publishers did not design materials for all school subjects and/or grade levels under investigation (see also Section 5.1), the statistical analysis did not allow for differentiation between publishers. To account for 1) potential differences in stylistic preferences between textbooks from different publishers and 2) correlations between texts for the two sub-domains of geography (which were selected from the same geography textbook, see also Section 5.1), *Textbook* was modeled as a random factor. Likelihood ratio tests were computed in order to assess which models fitted the data best. In the next section, we discuss the results of the best fitting models. See Appendix D for an overview of all results.

6 Results

In this section, we sketch a general picture of the occurrence of voice elements in our corpus (Section 6.1), before presenting the results of our statistical analyses (Section 6.2).

6.1 Overall occurrence of voice elements

Figure 1 shows the number of texts in which the different types of voice elements were found. It demonstrates that *you* is by far the most commonly used element in the corpus (451/1055 texts). In the remaining analyses, encouragements were grouped with imperatives, because of their low number of occurrence (19). Figure 2 indicates that more than half of the texts exhibit one up to three types of elements (608).

Only a small number of texts contain four or more types of voice elements (32), with two texts incorporating all six types. For example, in (11), we find the pronouns *you* and *we*, a question (*Did you know that everything you eat comes from plants?*), an imperative (*Just think of*), an exclamation (*And a cow eats plants!*), and an evaluation of to-be-learned information (*something special*).
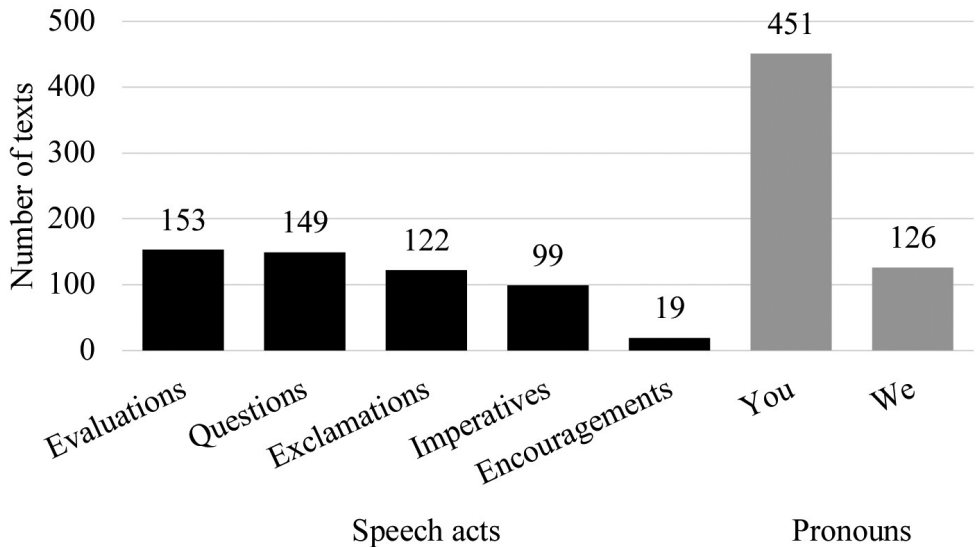
Figure 1 Number of texts in which the voice elements were found (N = 1055)
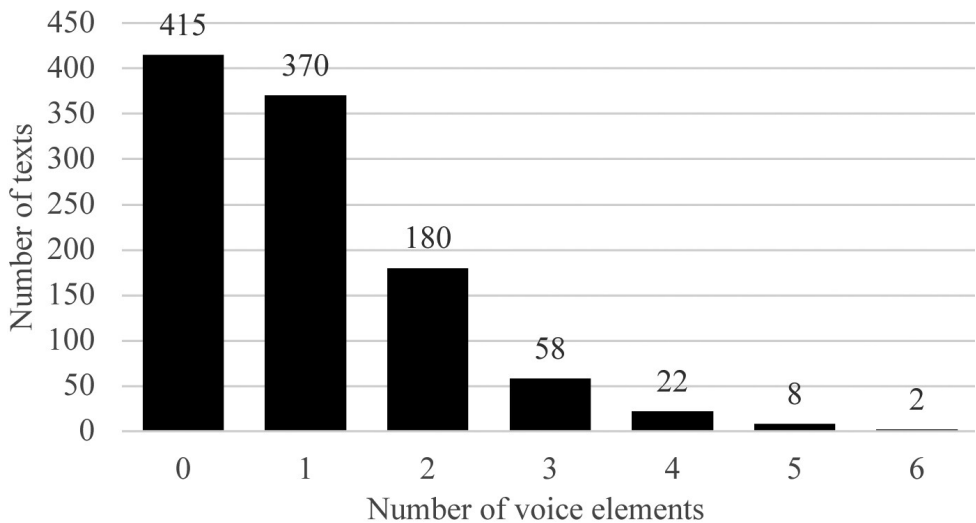
Figure 2 Number of texts per sum of different voice elements present in the text (N = 1055)

(11) *Did you know that everything you eat comes from plants? Just think of* strawberry jam and pasta, which is made of grain. Even when *you* eat meat, *you* actually eat plants. Because meat comes from a cow, for example. *And a cow eats plants!* Without plants, people and animals would have no food. There is *something special* going on with plants: they produce their own food. They do this in their leaves. Just like people and animals, plants are composed of very small living parts. *We* call them cells. *You* can compare them with bricks that form a house together.
(*Wijzer! Natuur & Techniek*, biology grade 5, p. 54)

6.2 Statistical analyses

First, we analyzed whether the distribution of educational texts with one or more types of voice elements (640/1055 texts) over the corpus was influenced by the fixed factors *Subject*, *Level*, and/or their interaction. For this analysis, which included all six types of voice elements, the best fitting model was the model in which only *Subject* was entered as a fixed factor ($\chi^2(3) = 21.83$, $p < .001$). A post hoc Tukey pairwise comparison test revealed that, in general, voice elements are less present in history texts than in texts for biology (*OR* = 3.96, *SE* = 1.18, $z = 4.64$, $p < .001$), physical geography (*OR* = 3.75, *SE* = 1.09, $z = 4.56$, $p < .001$), and human geography (*OR* = 2.60, *SE* = 0.75, $z = 3.35$, $p = .005$), which is visualized in Figure 3.

Subsequently, we clustered the six types of voice elements into two categories, one for speech acts (questions, imperatives, exclamations, evaluations) and one for pronouns (*you*, *we*). For speech acts, the model in which *Subject* and *Level*, but not their interaction, were entered as fixed factors fitted the data best ($\chi^2(3) = 8.58$, $p = .003$). A post hoc Tukey test revealed that speech acts are less frequent in history texts than in physical geography texts (*OR* = 2.77, *SE* = 0.80, $z = 3.54$, $p = .01$), and that they are more frequent in grade 5 texts than in grade 8 texts (*OR* = 2.17, *SE* = 0.52, $z = 3.24$, $p = .026$).

For pronouns, the best fitting model was the model in which only *Subject* was entered as a fixed factor ($\chi^2(3) = 31.22$, $p < .001$). A post hoc Tukey test showed the same pattern as for the analysis with all voice elements, namely that pronouns are less present in history texts than in texts for biology (*OR* = 5.65, *SE* = 1.65, $z = 5.92$, $p < .001$), physical geography (*OR* = 4.72, *SE* = 1.35, $z = 5.44$, $p < .001$), and human geography (*OR* = 2.89, *SE* = 0.82, $z = 3.74$, $p = .001$). The results for speech acts and pronouns are visualized in Figures 4 and 5.

Finally, we analyzed whether the general patterns persisted per voice element type. Below, we cluster the results by category (speech acts vs. pronouns). No interaction effects were found.
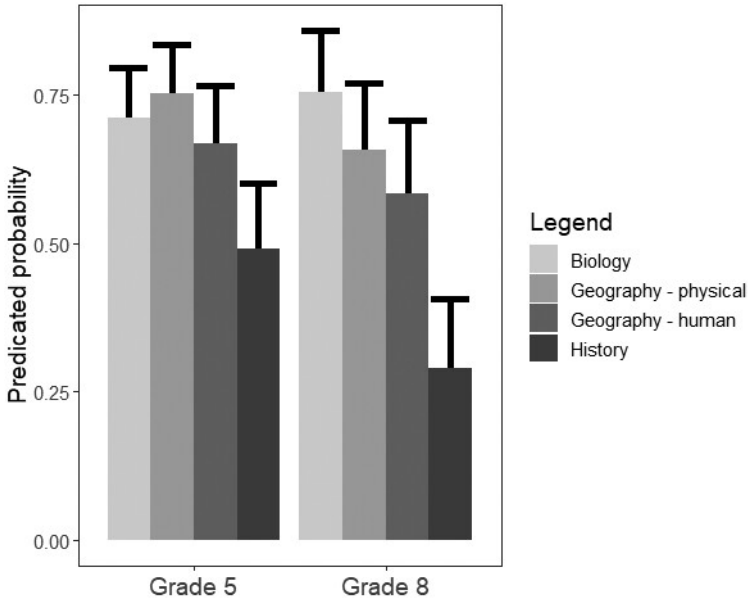
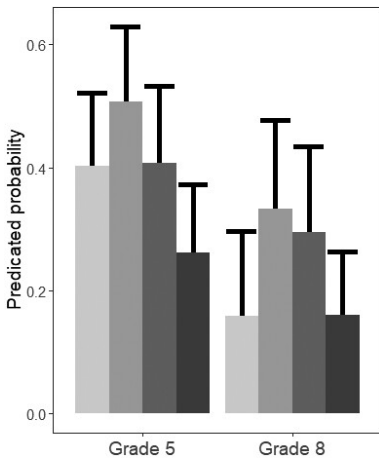**Figure 3** Predicted probability for texts with one or more voice elements

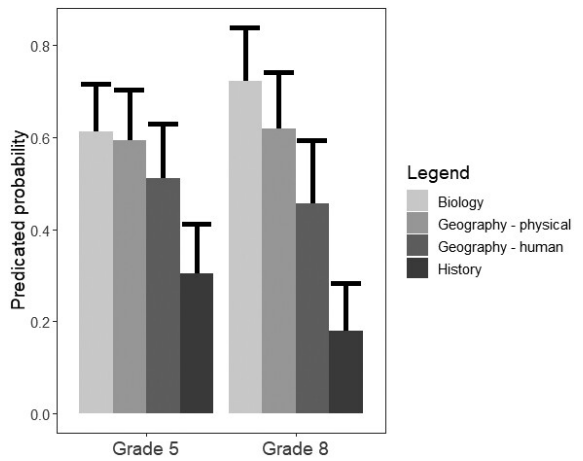**Figure 4** Predicted probability for speech acts

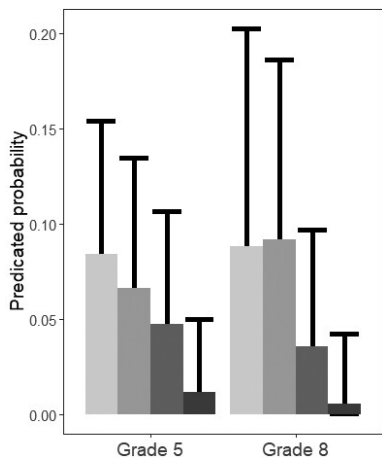**Figure 5** Predicted probability for pronouns

Figure 6 Predicted probability for imperatives
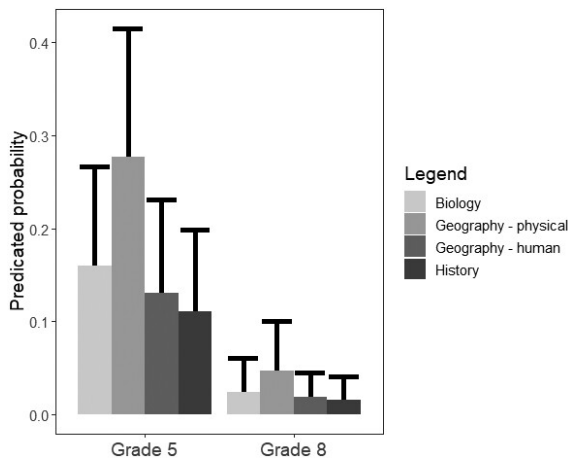
Figure 7 Predicted probability for exclamations

6.2.1 *Speech acts*

For questions and evaluations, no model significantly improved the base model. This implies a diversion from the general pattern for speech acts and shows that the distribution of these two types of elements is independent of *Subject* and *Level*.

For imperatives, the model in which *Subject* was entered as a fixed factor was the best fitting model ($\chi^2(3)$ = 19.49, p < .001). A post hoc Tukey test showed that imperatives are less frequent in history texts than in texts for biology (OR = 10.46, SE = 6.99, z = 3.51, p = .003) and physical geography (OR = 9.42, SE = 6.28, z = 3.37, p = .004).

For exclamations, the model in which both *Subject* and *Level* were entered as fixed factors was the best fitting model ($\chi^2(1)$ = 18.91, p = .006). A post hoc Tukey test showed that exclamations are less frequent in texts for human geography than in texts for physical geography (OR = 2.53, SE = 0.73, z = 3.22, p = .029), and that exclamations are more common in grade 5 texts than in grade 8 texts (OR = 7.70, SE = 3.28, z = 4.80, p < .001). The results for imperatives and exclamations are displayed in Figures 6 and 7.

6.2.2 *Pronouns*

Corresponding to the general patterns for pronouns, the best fitting model for *you* and *we* was the model in which only *Subject* was entered as a fixed factor (*you*: $\chi^2(3)$ = 27.46, p < .001; *we*: $\chi^2(3)$ = 9.23, p = .026). A post hoc Tukey test revealed that *you* is less present in history texts than in texts for biology (OR = 5.85, SE = 1.91, z = 5.41, p < .001), physical geography (OR = 4.97, SE = 1.59, z = 5.03, p < .001), and human geography (OR = 3.23, SE = 1.03, z = 3.68, p = .001), which is visualized in Figure 8. In addition, although a main effect of *Subject* was found for *we*, a post hoc Tukey test did not reveal any significant pairwise comparisons. Nonetheless, as Figure 9 shows, the pattern was quite similar to that of *you*.
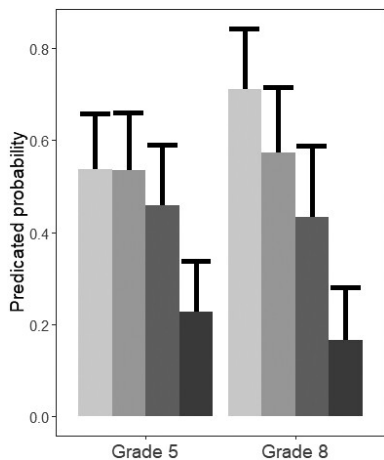
Figure 8 Predicted probability for *you*
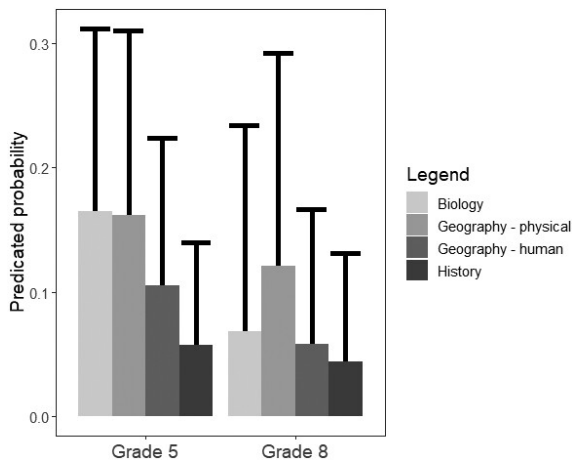
Figure 9 Predicted probability for *we*

7 Discussion and conclusion

In this study, we focused on voice elements: textual elements that the author of an educational text uses to interact with students. The inclusion of voice elements in educational texts seems to be an attractive strategy to enhance students' level of engagement, which is considered an important factor in text comprehension and learning from educational texts: students who are engaged in the active processing of educational texts are likely to be more attentive to the to-be-learned information represented in these texts, which stimulates deeper understanding and better learning of this information (cf. Beck et al., 1995; Brozo et al., 2007; Guthrie & Wigfield, 2000; Hidi, 2001; Sadoski, 2001; Schraw & Lehman, 2001). As such, voice elements can support the transfer of educational content.

By means of a quantitative corpus-based analysis, we investigated how and when voice elements are currently being used in Dutch educational texts, hypothesizing that the distribution of voice elements would be influenced by school subject (Hypothesis 1: HI < GH < GP = BI) and grade level (Hypothesis 2A: grade 5 > 8; Hypothesis 2B: grade 5 < 8). The significant pairwise patterns are summarized in Table 3.

Table 3 shows robust effects for the distribution of voice elements over school subjects. For both speech acts and pronouns, the effects generally support Hypothesis 1: most voice elements are less frequent in history texts compared to biology and geography texts. These findings suggest that educational publishers differentiate their use of voice elements over school subjects, taking the relatability of the educational content into account. Nevertheless, our distinction between physical and human geography texts appears to be less relevant than expected: rather than occupying an intermediate position

Table 3 Summary of the significant pairwise patterns

Voice elements	Main effect of subject	Main effect of level
All	HI < GH = GP = BI	
Speech acts	*HI < GP*	*5 > 8*
Questions		
Imperatives	HI < GP = BI	
Exclamations	GH < GP	5 > 8
Evaluations		
Pronouns	*HI < GH = GP = BI*	
You	HI < GH = GP = BI	
We	-*	

* Although a main effect of *Subject* was found, none of the pairwise comparisons reached significance.

between history texts on the one hand and biology and physical geography texts on the other hand, human geography texts tend to side with the latter two school subjects.

The lack of significant patterns in the third column of Table 3 signals that most voice elements are equally distributed over grade levels. Although an overall effect was found for the category of speech acts, this effect appears to be caused by only one element, namely exclamations. This indicates that for grade levels, neither Hypothesis 2A nor Hypothesis 2B is accepted: 1) voice elements are not used more frequently in secondary education texts to bridge an increasing gap between the educational content and students (2B "bridging student-content gap"), and 2) only limited evidence was found for the hypothesis that educational publishers strive for a progression from relatively simple to more challenging educational texts in their use of voice elements (2A "progression simple-challenging").

In what follows, we aim at clarifying the similarities and variations in the distribution of voice elements over the corpus, discussing the general lack of differences between grade levels (Section 7.1), and seeking explanations for the variations in the distribution over school subjects (Section 7.2). Finally, we present directions for future research (Section 7.3).

7.1 Differentiation over grade levels

The occurrence of the different voice elements is highly comparable across grade levels, as only exclamations are more frequent in grade 5 texts than in grade 8 texts. This finding supports Hypothesis 2A, which predicted a decrease in the use of voice elements as texts at higher grades should display a progression from relatively simple to more

challenging (cf. Brabham & Villaume, 2002; Shanahan et al., 2012; Snow, 2002). There may be two reasons why such an effect is found for exclamations. First, exclamations are an interactive means to express emotional and attitudinal aspects of communication that are easily recognizable in face-to-face interaction (e.g., tone of voice), but remain generally unobservable in written language (Bonvillain, 2020; Nystrand, 1986). Authors can use exclamations to stress content they consider important or remarkable, thus directing students' attention to specific parts of the educational text. They may believe that grade 5 students are less able to identify important and remarkable content than grade 8 students. Second, exclamations are generally considered less appropriate in an expository context (Clark & Pointon, 2016). Therefore, they tend to have a more informal status than other types of voice elements. This informality helps to put into words the amazement children experience when discovering new things in their world. As such, publishers may consider exclamations appropriate to enthuse children in grade 5, while they may choose to avoid using this kind of language in higher grades, because it may be experienced as childish. Of course, the two reasons may also be related.

Hypothesis 2A is not supported by the findings for the other types of voice elements. There may be several reasons why. For instance, it could be the case that publishers strategically consider the distribution of voice elements over educational texts, but that a decrease in use to enhance the complexity of texts (5 > 8, Hypothesis 2A) is cancelled out by an increase in use to bridge the increasing student-text gap (5 < 8, Hypothesis 2B). Alternatively, it could be the case that publishers are not as strategic in this respect as we expected them to be. That is, they might actually not pursue a deliberate distribution strategy with respect to the use of voice elements in educational texts, resulting in an accidental equal distribution of voice elements over texts for grade 5 and grade 8. However, such explanations cannot be substantiated on the basis of a corpus-based analysis alone. To this end, it would be relevant to discuss the current findings with publishers (see also Section 7.3).

Finally, as none of the findings of our corpus-based analysis support Hypothesis 2B, which predicted an increase in the use of voice elements to bridge an increasing student-content gap, one could wonder to what extent there actually is variation in the student-content distance between different grade levels. In absolute terms, the educational content tends to be more complex in secondary education texts than in primary education texts (Committee Meijerink, 2009; Hidi 2000). However, in relative terms, the complexity of the educational content seems to be 'growing' with students; concurrent with their expanding background knowledge, the to-be-learned information gradually becomes more challenging, while the student-content distance remains unchanged. In this respect, it may be less necessary to increase the number of voice elements when shifting from primary to secondary education texts.

7.2 Differentiation over school subjects

The findings for school subjects generally support Hypothesis 1 (HI < GH < GP = BI), for both speech acts and pronouns. While *you* and *we* can be used in all school subjects relatively easily, they were found to be less frequent in history texts. This can be explained by the fact that in these texts, historical figures act as identifiable go-betweens: their presence makes it easier for students to view the educational content from a different perspective, making the to-be-learned information easily imaginable and relatable, and allowing the text's author to stay in the background (cf. Hidi, 2001; Kuijpers, 2014). By contrast, in biology and geography texts, it seems to be more essential for the author to step in as a mediator and directly address students to involve them with the educational content (cf. Nolen, 1995; Sangers, 2022; Sangers et al., 2021). Similarly, for imperatives, the need to instruct students to react to the educational content (e.g., by active thinking) seems to be less urgent in history texts than in biology and physical geography texts.

The lack of difference in the use of questions across school subjects may be explained by the fact that questions are a neutral and simple didactic tool to arouse students' curiosity to discover new information; questions can be used rather effortlessly to provide rhetoric structures to educational texts and/or lessons, stimulating inquiry-based learning (cf. Caram & Davis, 2005; Chaudhari, 1974). The explanation for the lack of difference in the use of evaluations, however, is less evident. For instance, it is not a matter of a limited number of instances (153/1055 texts). Nonetheless, we noticed that evaluations can be quite subtle, especially when formulated as adjectives (e.g., *He was a* smart *commander*; *Horses, donkeys, and zebras yield* beautiful *crossings*). When writing educational texts, and particularly when attempting to write them in a vivid rather than dull way, it might be difficult to leave such evaluations out. Alternatively, our selection criteria and 0-bias might have been too strict to prompt pronounced variation in the distribution of evaluations.

7.3 Suggestions for future research

The present research has uncovered similarities as well as variations in the distribution of voice elements over Dutch educational texts. In future research, it would be worthwhile to discuss our findings with educational publishers. This would allow us to find out whether our findings arise from explicit design principles set by publishers or merely from individual authors' intuitions. To what extent do publishers deliberately vary their use of voice elements over school subjects, grade levels and/or educational levels? And to what degree are such design principles agreed upon between different authors or design teams working for the same publishing company? Earlier research has shown that interviews can be fruitful in discovering what publishers consider important textual elements, and how they adapt their design principles accordingly (Land et al., 2002).

Future research should also focus on the actual effects that voice elements have on students' engagement and comprehension of educational texts. To what extent do voice elements affect the distance between students and the educational content? To what extent do they enhance students' motivation and reading processes? Do different (combinations of) types of voice elements lead to different effects? And to what extent do these effects fit publishers' design principles, intuitions, and current practices? Although the findings of Beck et al. (1995) provide a first indication that voice leads to beneficial effects on text comprehension (see also Section 1), it remains unclear what the exact effects of including voice elements in educational texts are, especially when focusing on elements that directly relate to the educational text's author.

Furthermore, there are some limitations to our study that give rise to new directions for future research. First, we focused on biology, geography, and history texts for grade 5 and grade 8. In future research, it would be interesting to expand the current research by examining the distribution of voice elements over additional school subjects and grade levels, and also make international comparisons to educational texts for other countries than the Netherlands. In addition, it would be worthwhile to explore the use of automatic analyses; not only to expedite the method of analysis but also to more conveniently allow for inclusion of the relative distribution of the different types of voice elements over educational texts, which was left out of consideration in this study. Finally, besides voice elements, publishers can use other linguistic strategies to make their educational texts more engaging and comprehensible. These strategies provide fruitful directions for further study. One such direction is the narrative component of Beck et al.'s (1995) interpretation of voice, which was left aside in the present study, since our aim was to chart the distribution of textual elements that the author of an educational text uses to interact with students. Nevertheless, other voices can be incorporated in educational texts, such as those of narrative characters. The addition of such characters, as well as other narrative elements, to educational texts could be another means to provide students with proper, relatable contexts for the information they need to learn, thereby enhancing their reading processes (cf. Sangers, 2022; Sangers et al., 2021).

It is hard to overestimate the importance of well-designed educational texts that provide relevant content in an engaging and comprehensible way. The inclusion of voice in educational texts could serve an important role in achieving this goal. The current study has specified Beck et al.'s (1995) notion of voice by defining it solely in terms of textual elements that directly relate to the author of an educational text. Our study has shown that author-initiated voice elements are quite extensively used in Dutch educational texts, and that their use is influenced by school subject. This suggests that authors have intuitions about when voice elements are particularly helpful. A next step is to examine whether such intuitions are integrated into actual design principles concerning the use of voice elements in educational texts. In addition, having charted the current practices of using voice elements in educational texts, this study has provided an essential step to further investigate the promise of voice elements for designing better educational texts.

Notes

1   The operations used for the coherent versions were clarifying, elaborating, explaining, providing motivation for important information, and making connections explicit (Beck et al., 1995).

2   For instance, "each point at which the skilled writer chooses one example rather than another, one term rather than another, certain comparisons rather than others, etc., is ultimately arbitrated not only by what the writer has to say but also by the needs of his or her readers to understand." (Nystrand, 1986, p. 36).

3   Throughout the paper, we represent examples from Dutch textbooks by their English translations, in which the relevant voice elements are italicized. For an overview of the Dutch examples, see Appendix A.

4   The term "publishers" is used as an umbrella term here, as it refers to both the editors and policy makers that work at a publishing company, as well as the actual educational content writers.

5   The Dutch system for secondary education is divided into three educational levels, ranging from theoretical to vocational training: pre-university education (Dutch *vwo*), senior general education (Dutch *havo*), and pre-vocational education (Dutch *vmbo*). Dutch primary education does not make this distinction, which is why the students in grade 5 form a diverse group with different intellectual levels.

6   The Cohen's kappa score for encouragements to imagine is moderate because this element is only present in 19/1055 texts (1.8% of the corpus), see Section 6.1. When combining encouragements with imperatives, the inter-annotator agreement reaches $\kappa = .85$ and 97%.

7   In Dutch primary education, biology is called "nature and technique" (*Natuur en Techniek*), combining biological and physical phenomena. Since physical phenomena are not part of the biology curriculum in pre-university education (instead they are discussed in a separate school subject called "physics"), we excluded chapters that focused on physics.

8   The asterisk indicates the model that was proven to be the best fitting model.

References

Andeweg, S., Hendrix, R., Van 't Hoff, V., & De Hoop, H. (2013). En dan vermoord je hem …: Over de invloed van grammaticale persoon op identificatie [And then you murder him …: On the influence of grammatical person on identification]. *Neerlandistiek.nl*, *13*(1), 1–19.

Austin, J.L. (1962). *How to do things with words*. Oxford University Press.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Beck, I.L., McKeown, M.G., Sinatra, G.M., & Loxterman, J.A. (1991). Revising social studies text from a text-processing perspective: Evidence of improved comprehensibility. *Reading Research Quarterly*, 26(3), 251–276. https://doi.org/10.2307/747763

Beck, I.L., McKeown, M.G., & Worthy, J. (1995). Giving a text voice can improve students' understanding. *Reading Research Quarterly*, 30(2), 220–238. https://doi.org/10.2307/748033

Bonvillain, N. (2020). *Language, culture, & communication: The meaning of messages* (8th ed.). The Rowman & Littlefield Publishing Group.

Brabham, E.G., & Villaume, S.K. (2002). Leveled text: The good and the bad news. *The Reading Teacher*, 55(5), 438–441.

Brozo, W.G., Shiel, G., & Topping, K. (2007). Engagement in reading: Lessons learned from three PISA countries. *Journal of Adolescent & Adult Literacy*, 51(4), 304–315. https://doi.org/10.1598/JAAL.51.4.2

Caram, C.A., & Davis, P.B. (2005). Inviting student engagement with questioning. *Kappa Delta Pi Record*, 42(1), 19–23. https://doi.org/10.1080/00228958.2005.10532080

Chaudhari, U.S. (1974). The role of questions in thinking and learning from text: A research perspective. *Educational Technology*, 14(1), 7–11.

Clark, S., & Pointon, G. (2016). *The Routledge student guide to English usage: A guide to academic writing for students*. Routledge.

Conrad, S., & Biber, D. (2000). Adverbial marking of stance in speech and writing. In S. Hunston, & G. Thompson (Eds.), *Evaluation in text: Authorial stance and the construction of discourse* (pp. 56–73). Oxford University Press.

Committee Meijerink (2009). *Referentiekader taal en rekenen: De referentie-niveaus* [Framework of reference for language and arithmetic: The reference levels]. Doorlopende Leerlijnen Taal en Rekenen.

De Hoop, H., & S. Tarenskeen (2012). It's all about you in Dutch. *Journal of Pragmatics*, 88, 163–175. https://doi.org/10.1016/j.pragma.2014.07.001

Deringer, L., Gast, V., Haas, F., & Rudolf, O. (2015). Impersonal uses of the second person singular and generalized empathy: An exploratory corpus study of English, German and Russian. In L. Gardelle & S. Sorin (Eds.), *The pragmatics of personal pronouns* (pp. 311–334). John Benjamins.

Dood, C., Gubbels, J. & Segers, P.C.J. (2020). *PISA-2018 De verdieping: Leesplezier, zelfbeeld bij het lezen, leesgedrag en leesvaardigheid en de relatie daartussen* [PISA-2018: The deepening: Enjoyment of reading, self-perception when reading, reading behavior and reading comprehension and the relationship between them]. Expertisecentrum Nederlands.

Evers-Vermeul, J. & Holtermann, M. (2013). Doorlopende leerlijnen: implicaties voor leveling van leer- en examenteksten voor het middelbaar onderwijs [Continuous learning: implications for the leveling of learning and examination texts in secondary education]. *Tijdschrift voor Taalbeheersing*, 35(1), 1–24. https://doi.org/10.5117/TVT2013.1.EVER

Gast, V., Deringer, L., Haas, F., Rudolf, O. (2015). Impersonal uses of the second person singular: A pragmatic analysis of generalization and empathy effects, *Journal of Pragmatics*, 88, 148–162. https://doi.org/10.1016/j.pragma.2014.12.009

Gubbels, J., Netten, A., & Verhoeven, L. (2017). *Vijftien jaar leesprestaties in Nederland: PIRLS-2016* [Fifteen years of reading achievement in the Netherlands: PIRLS-2016]. Expertisecentrum Nederlands.

Gubbels, J., van Langen, A.M.L., Maassen, N.A.M., & Meelissen, M.R.M. (2019). *Resultaten PISA-2018 in vogelvlucht* [Results PISA-2018 in a bird's eye view]. Universiteit Twente.

Guthrie, J.T., & Wigfield, A. (2000). Engagement and motivation in reading. In M.L. Kamil, P.B. Mosenthal, P.D. Pearson, & R. Barr (Eds.), *Handbook of reading research: Volume III* (pp. 403–422). Routledge Taylor & Francis Group.

Hidi, S. (2000). An interest researcher's perspective on the effects of extrinsic and intrinsic factors on motivation. In C. Sansone, & J.M. Harackiewicz (Eds.), *Intrinsinc motivation: controversies and new directions* (pp. 309–339), Academic Press.

Hidi, S. (2001). Interest, reading, and learning: Theoretical and practical considerations. *Educational Psychology Review*, 13(3), 191–209. https://doi.org/10.1023/A:1016667621114

Hogeweg, L., & De Hoop, H. (2015). Introduction: The flexibility of pronoun reference in context. *Journal of Pragmatics*, 88, 133–136. https://doi.org/10.1016/j.pragma.2015.06.009

Houtkoop, H., & Koole, T. (2000). *Taal in actie: Hoe mensen communiceren met taal* [Language in action: How people communicate through language]. Uitgeverij Coutinho.

Inspectorate of Education (2017). *De staat van de leerling* [The state of the student].

Inspectorate of Education (2020). *De staat van het onderwijs 2020* [The state of education].

Inspectorate of Education (2021). *De staat van het onderwijs 2021* [The state of education].

Kuijpers, M. (2014). Absorbing stories: The effects of textual devices on absorption and evaluative responses. [Doctoral dissertation, Utrecht University].

Land, J., Sanders, T., Lentz, L., & Van den Bergh, H. (2002). *Tekstbegrip en tekstwaardering op het vmbo: Welke tekstkenmerken dragen bij aan de kwaliteit van studieboekteksten?* [Text comprehension and appreciation in lower secondary professional education: Which textual features contribute to the quality of educational texts?]. Stichting Lezen; Utrecht University.

Landis, J.R., & Koch, G.G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33(1), 159–174. https://doi.org/10.2307/2529310

Lenth, R. (2019). Emmeans: Estimated marginal means, aka least-squares means (Version 1.4.1) [Computer software]. https://cran.r-project.org/web/packages/emmeans

McKeown, M. G, Beck, I. L, Sinatra, G.M., Loxterman, J.A. (1992). The contribution of prior knowledge and coherent text to comprehension. *Reading Research Quarterly*, 27(1), 79–93. https://doi.org/10.2307/747834

Neuendorf, K.A. (2002). *The content analysis handbook*. Sage Publications.

Nolen, S.B. (1995). Effects of a visible author in statistical texts. *Journal of Educational Psychology*, 87(1), 47–65. https://doi.org/10.1037/0022-0663.87.1.47

Nystrand, M. (1986). *The structure of written communication: Studies in reciprocity between writers and readers*. Academic Press.

Orvell, A., Kross, E., & Gelman, S.A. (2020). "You" speaks to me: Effects of generic-you in creating resonance between people and ideas. *Psychological and Cognitive Sciences*, 117(49), 31038-31045. https://doi.org/10.1073/pnas.2010939117

Pander Maat, H. (2002). *Tekstanalyse: Wat teksten tot teksten maakt* [Tekst analysis: what makes texts texts]. Uitgeverij Coutinho.

R Core Team (2019). R: A language and environment for statistical computing. R foundation for statistical computing (Version 3.6.1) [Computer software]. https://www.r-project.org/

Sadoski, M. (2001). Resolving the effects of concreteness on interest, comprehension, and learning important ideas from text. *Educational Psychology Review*, 13(3), 263–281. https://doi.org/10.1023/A:1016675822931

Sangers, N.L. (2022). *Vivid educational texts: Engaging students via narrative and voice elements* [Doctoral dissertation, Utrecht University]. LOT. https://dx.medra.org/10.48273/LOT0618

Sangers, N.L., Evers-Vermeul, J., Sanders, T.J.M., & Hoeken, H. (2020). Vivid elements in Dutch educational texts. *Narrative Inquiry*, 30(1), 185–209. https://doi.org/10.1075/ni.18090.san

Sangers, N.L., Evers-Vermeul, J., Sanders, T.J.M., & Hoeken, H. (2021). Narrative elements in expository texts: A corpus study of educational textbooks. *Dialogue & Discourse*, 12(2), 115–144. https://doi.org/10.5210/dad.2021.204

Schleppegrell, M.J. (2004). *The language of schooling: A functional linguistics perspective*. Lawrence Erlbaum.

Schraw, G., & Lehman, S. (2001). Situational interest: A review of the literature and directions for future research. *Educational Psychology Review*, 13(1), 23–52. https://doi.org/10.1023/A:1009004801455

Searle, J.R. (1969). *Speech acts: An essay in the philosophy of language*. Cambridge University Press.

Searle, J.R. (1975). Indirect speech acts. In P. Cole, & J.L. Morgan (Eds.), *Syntax and semantics. Vol III: Speech acts* (pp. 59–82). Academic Press.

Shanahan, T., Fisher, D., & Frey, N. (2012). The challenge of challenging text. In M. Scherer (Ed.), *On developing readers: Readings from educational leadership* (pp. 100–109). ASCD.

Smolkin, L.B., McTigue, E.M., Donovan, C.A., & Coleman, J.M. (2008). Explanation in science trade books recommended for use with elementary students. *Science Education*, 93(4), 587–610. https://doi.org/10.1002/sce.20313

Snow, C. (2002). *Reading for understanding: Toward an R&D program in reading comprehension*. CA: RAND.

Toolan, M. (2001). *Narrative: A critical linguistic introduction* (2nd ed.). Routledge.

Vološinov, V.N. (1973). *Marxism and the philosophy of language* (L. Matejka & I.R. Titunik, Trans.). Seminar Press.

Wechsler, S. (2010) What 'you' and 'I' mean to each other: person indexicals, self-ascription, and theory of mind. *Language,* 86(2), 332–365. https://doi.org/10.1353/lan.0.0220

Wickham, H. (2016). Ggplot2: Elegant graphics for data analysis [Computer software]. https://ggplot2.tidyverse.org

Wickham, H., & Miller, E. (2019). Haven: Import and export "SPSS", "Stata" and "SAS" Files (Version 2.1.1) [Computer software]. https://CRAN.R-project.org/package = haven

Woldhuis, E., Rodenboog, M., Heijnen, M, & Fisser, P. (2018). *Leerlingenmonitor 17/18: Leermiddelen in het po en vo* [Studentmonitor 17/18: Instructional equipment in primary and secondary education]. SLO.

Appendix A – Original Dutch examples

(1) Vind jij het belangrijk dat je kleren er leuk uitzien of lekker zitten? Vroeger vonden mensen kleren vooral belangrijk om hun lichaam te beschermen tegen het weer of tegen schrammen van takken. De kleding moest dus stevig en lekker warm zijn.

(2) Zet een helm op en probeer ook een stofbril te vinden. Ga een minuut omgekeerd aan het klimrek hangen. Richt af en toe je bovenlichaam op.

(3) 's Ochtends onder de douche met een krachtige waterstraal is heerlijk! Maar stel je voor dat er zand in het water zit: je huid wordt gezandstraald, en dat is minder prettig. Laat dat nou precies datgene zijn wat rivieren doen met gesteente …

(4) Sommige zaadplanten kunnen zich ook voortplanten zonder zaden! Hoe ze dat doen? Door een nieuw plantje te laten groeien uit een stengel, een wortel, een knol of een bol.

(5) Jij hoort erbij: de Europese Unie. Wel 28 verschillende landen volgen dezelfde regels en wetten. Ze werken samen op economisch en politiek gebied. Sommige landen zijn rijk, andere wat minder. De mensen spreken er verschillende talen, hebben andere gebruiken en toch gaat het meestal goed. Bijzonder hoor, die Europese Unie!

(6) Vanuit de ruimte moet je goed kijken om het kleine Nederland te zien. Maar als je weet waar de Noordzee ligt, heb je het snel gevonden. De Noordzee bepaalt ons klimaat. We hebben hier een gematigd zeeklimaat.

(7) Door de schuine stand van de aardas komt de zon op het noordelijk halfrond in de zomer hoog boven de horizon. Helemaal in het noorden gaat hij 's nachts zelfs niet onder. Van 20 mei tot 22 juli is het er dag en nacht licht. Dat noemen we de middernachtzon. In de winter wordt het in dit gebied juist helemaal niet licht. Dat heet de poolnacht.

(8) De Europese Unie heeft, net als Nederland, een eigen regering: de Europese Commissie (EC). De Europese Commissie zit in Brussel. In de EC zitten 27 commissarissen, een soort ministers. Deze worden benoemd door de regering van hun eigen land. Iedere lidstaat levert één commissaris, hoe groot of klein het land ook is.

(9) 'Guten Tag! Ik ben Matthias Sammer. Ik woon in Berlijn en vind het leuk om jouw gids te zijn bij de verkenning van mijn land.'

(10) Door niezen, hoesten, zoenen of door elkaars bestek te gebruiken, kan de ziekte [AIDS] niet worden overgebracht. Je kunt dus zonder gevaar op een normale manier met aidspatiënten of seropositieve mensen blijven omgaan.

(11) Wist jij dat alles wat je eet, van planten komt? Denk maar aan aardbeienjam en aan pasta, dat van graan wordt gemaakt. Ook als je vlees eet, eet je eigenlijk planten. Want vlees komt bijvoorbeeld van een koe. En een koe eet planten! Zonder planten zouden mensen en dieren geen voedsel hebben. Met planten is iets bijzonders aan de hand: ze maken hun eigen voedsel. Dat doen ze in hun bladeren. Net als mensen en dieren zijn planten opgebouwd uit heel kleine levende deeltjes. We noemen ze cellen. Je kunt ze vergelijken met bakstenen die samen een huis vormen.

Appendix B – Materials

Biology: grade 5

Ottenheim, M., & Tromp, R. (Eds.) (2011). *Natuniek: Natuur en Techniek voor het basisonderwijs groep 7 leerlingenboek*. ThiemeMeulenhoff.

Siemensma, F. (2014). *Wijzer! Natuur & Techniek: Leerwerkboek groep 7*. Noordhoff Uitgevers.

Talsma, A., & Vogelesang, L. (Eds.) (n.d.). *Natuurzaken: Werkboek jaargroep 7* (4th ed.). Uitgeverij Zwijsen.

Van Riel, M., & Soet, L. (Eds.). (2012). *Argus Clou professor in alles: Natuur en Techniek groep 7 lesboek*. Malmberg.

Wiechers, C. (Ed.) (2014). *Binnenstebuiten: Natuur en Techniek bronnenboek groep 7*. Blink Educatie.

Biology: grade 8

Akkerman, T. (Ed.) (2013). *Nectar: 2–3 vwo leerboek* (4th ed.). Noordhoff Uitgevers.

Bos, A., Kalverda, O., Passier, R., Rawee, H., Smale, R., Smits, G., & Waas, B. (2015). *Biologie voor jou: Handboek 2a vwo/gymnasium* (7th ed.). Malmberg.

Geography: grade 5

Bakker, A. (Ed.) (2012). *De blauwe planeet: Aardrijkskunde voor het basisonderwijs*. ThiemeMeulenhoff.

Huisman, A. (Ed.) (2012). *Argus Clou professor in alles: Aardrijkskunde groep 7 lesboek*. Malmberg.

Siemensma, F. (2015). *Wijzer! Aardrijkskunde: Leerwerkboek groep 7*. Noordhoff Uitgevers.

Talsma, A. (Ed.) (2014). *Wereldzaken: Werkboek jaargroep 7* (3rd ed.). Uitgeverij Zwijsen.

Van Ooijen, M. (Ed.) (2014). *Grenzeloos: Aardrijkskunde bronnengroep groep 7*. Blink Educatie.

Geography: grade 8

Ariaens, D., ten Brinke, W., de Jong, C., & Padmos, J.H.A. (Eds.) (2016). *De Geo: Aardrijkskunde voor de onderbouw 2 vwo* (9th ed.). ThiemeMeulenhoff.

Van de Ven, M. (Ed.) (2018). *De wereld van: Aardrijkskunde voor de onderbouw leeropdrachtenboek 2VG*. Malmberg.

Van den Berg, G. (Ed.) (2014). *BuiteNLand 2 vwo* (3rd ed.). Noordhoff Uitgevers.

History: grade 5

Kruis, M. (2014). *Wijzer! Geschiedenis: Leerwerkboek Groep 7*. Noordhoff Uitgevers.

Können, A. (Ed.) (2012). *Argus Clou professor in alles: Geschiedenis groep 7 lesboek*. Malmberg.

Nijman, J., & Roest, H. (2011). *Speurtocht 7: Geschiedenis voor het basisonderwijs: Leerlingenboek* (2nd ed.). ThiemeMeulenhoff.

Van de Mortel, M., van den Oever, M., Vermeer, H., & Vogelesang, L. (n.d.) *Tijdzaken: Werkboek jaargroep 7* (4th ed.). Uitgeverij Zwijsen.

Wiechers, C. (Ed.) (2014). *Eigentijds: Geschiedenis bronnenboek groep 7*. Blink Educatie.

History: grade 8

Salemink, L., & Venner, J. (Eds.) (2010). *Feniks: Geschiedenis voor de onderbouw leesboek 2 vwo*. ThiemeMeulenhoff.

Schrover, W., & Tadema, J. (Eds.) (2015). *Memo: Geschiedenis voor de onderbouw 2 vwo handboek* (4th ed.). Malmberg.

Van der Geugten, T., & Verkuil, D. (Eds.) (2013). *Geschiedenis werkplaats: 2 vwo informatieboek*. (2nd ed.). Noordhoff Uitgevers.

Appendix C – Selection of themes

For history, we selected chapters that discussed the time period of Dutch stadtholder William of Orange, who led the Dutch Revolt against Spain during the start of the Eighty Years' war (1568–1648). One grade 5 textbook did not discuss this theme, mentioning only events after 1900. For this textbook, we selected a chapter that also discussed a war, namely one on the Cold War.

For biology, the reproduction of humans, animals, and plants was chosen as the overlapping theme. In grade 5, these kinds of reproduction are discussed together, while grade

8 textbooks mainly focus on the reproduction of humans. One selected grade 5 textbook did not include information on reproduction. For this textbook, we selected a chapter on eating habits of animals and plants, which had the highest number of keywords in common with the already selected chapters on the reproduction theme.[7]

For geography, chapters were matched within grade level only, since it turned out to be unfeasible to select thematically overlapping chapters between grade levels. For physical geography, the grade 5 chapters were matched by their discussion of different landscapes, and the grade 8 chapters by their focus on characteristics of the earth. For human geography, the grade 5 chapters concentrated on the European Union, while the grade 8 chapters focused on demographic notions such as "birth rate" and "immigration". Although the distinction between chapters on physical versus human geography topics was generally straightforward, one grade 5 textbook paid equal attention to both sub-domains in all its chapters. We selected two chapters that shared the most keywords with the other selected chapters, leading to the selection of a chapter on the climates and landscapes of Eastern Europe (physical geography), and a chapter on Europe, including discussions on the European Union (human geography).

Appendix D – Results

1 Generalized Linear Mixed Models[8]

Total	−2LL	$\Delta\chi^2$	Δdf	p
Model 0	1368.3			
*Model 1 (+SUBJECT)	1346.5	21.83	3	< .001
Model 2 (+LEVEL)	1343.1	3.35	1	.067
Model 3 (+SUBJECT:LEVEL)	1339.3	3.78	3	.029

Speech acts	−2LL	$\Delta\chi^2$	Δdf	p
Model 0	1295.3			
Model 1 (+SUBJECT)	1285.3	9.04	1	.029
*Model 2 (+LEVEL)	1276.8	8.58	3	.003
Model 3 (+SUBJECT:LEVEL)	1274.9	1.83	1	.609

Pronouns	−2LL	Δχ²	Δdf	p
Model 0	1387.3			
*Model 1 (+SUBJECT)	1356.1	31.22	1	< .001
Model 2 (+LEVEL)	1355.8	0.29	3	.591
Model 3 (+SUBJECT:LEVEL)	1351.0	4.81	1	.187

Questions	−2LL	Δχ²	Δdf	p
*Model 0 (base model)	813.4			
Model 1 (+SUBJECT)	809.7	3.62	3	.305
Model 2 (+LEVEL)	807.7	2.09	1	.149
Model 3 (+SUBJECT:LEVEL)	805.4	2.21	3	.530

Evaluations	−2LL	Δχ²	Δdf	p
*Model 0	852.8			
Model 1 (+SUBJECT)	847.2	5.52	3	.138
Model 2 (+LEVEL)	846.8	0.46	1	.500
Model 3 (+SUBJECT:LEVEL)	841.0	5.84	3	.120

Imperatives	−2LL	Δχ²	Δdf	p
Model 0	445.9			
*Model 1 (+SUBJECT)	426.4	19.49	3	< .001
Model 2 (+LEVEL)	426.4	0.002	1	.963
Model 3 (+SUBJECT:LEVEL)	425.3	1.14	3	.767

Exclamations	−2LL	Δχ²	Δdf	p
Model 0	689.6			
Model 1 (+SUBJECT)	677.3	12.33	3	.006
*Model 2 (+LEVEL)	658.4	18.91	1	< .001
Model 3 (+SUBJECT:LEVEL)	656.3	2.11	3	.550

You	−2LL	Δχ²	Δdf	p
Model 0	1342.3			
*Model 1 (+SUBJECT)	1314.8	27.46	3	< .001
Model 2 (+LEVEL)	1314.7	0.12	1	.729
Model 3 (+SUBJECT:LEVEL)	1311.1	3.55	3	.315

We	−2LL	Δχ²	Δdf	p
Model 0	730.8			
*Model 1 (+SUBJECT)	721.6	9.23	3	.026
Model 2 (+LEVEL)	720.0	1.59	1	.208
Model 3 (+SUBJECT:LEVEL)	719.3	0.72	3	.870

2 Predicted probability scores

Total	Probability	SE	LCL	UCL
Biology	0.73	0.04	0.61	0.82
Geography – physical	0.72	0.04	0.60	0.81
Geography – human	0.64	0.05	0.51	0.75
History	0.40	0.05	0.29	0.53

Speech acts		Probability	SE	LCL	UCL
Grade 5	Biology	0.37	0.05	0.24	0.52
	Geography – physical	0.51	0.06	0.36	0.66
	Geography – human	0.44	0.06	0.30	0.59
	History	0.27	0.05	0.17	0.41
Grade 8	Biology	0.21	0.05	0.11	0.37
	Geography – physical	0.33	0.05	0.20	0.49
	Geography – human	0.26	0.05	0.15	0.41
	History	0.15	0.03	0.08	0.26

Pronouns	Probability	SE	LCL	UCL
Biology	0.65	0.05	0.52	0.76
Geography – physical	0.61	0.05	0.48	0.72
Geography – human	0.49	0.05	0.37	0.61
History	0.25	0.04	0.17	0.35

Imperatives	Probability	SE	LCL	UCL
Biology	0.09	0.02	0.05	0.14
Geography – physical	0.08	0.02	0.05	0.13
Geography – human	0.04	0.01	0.02	0.08
History	0.01	0.01	0.003	0.03

Exclamations		Probability	SE	LCL	UCL
Grade 5	Biology	0.16	0.04	0.09	0.27
	Geography – physical	0.28	0.06	0.17	0.42
	Geography – human	0.13	0.04	0.07	0.23
	History	0.11	0.03	0.06	0.20
Grade 8	Biology	0.02	0.01	0.01	0.06
	Geography – physical	0.05	0.02	0.02	0.10
	Geography – human	0.02	0.01	0.01	0.05
	History	0.02	0.01	0.01	0.04

You	Probability	SE	LCL	UCL
Biology	0.59	0.06	0.48	0.70
Geography – physical	0.55	0.05	0.45	0.66
Geography – human	0.45	0.05	0.34	0.55
History	0.20	0.04	0.14	0.28

We	Probability	SE	LCL	UCL
Biology	0.13	0.04	0.07	0.24
Geography – physical	0.15	0.04	0.08	0.25
Geography – human	0.08	0.03	0.04	0.16
History	0.05	0.02	0.02	0.11

3 Post hoc Tukey scores*

*Note: BI = biology, GH = human geography, GP = physical geography, HI = history, 5 = grade 5, 8 = grade 8

Total

Contrasts			OR	SE	z	p
BI	/	GH	1.52	0.46	1.38	.510
BI	/	GP	1.06	0.32	0.18	.998
BI	/	HI	3.96	1.18	4.64	< .001
GH	/	GP	0.69	0.13	−1.91	.226
GH	/	HI	2.60	0.75	3.35	.005
GP	/	HI	3.75	1.09	4.56	< .001

Speech acts

Contrasts			OR	SE	z	p
5BI	/	5GH	0.76	0.23	−0.91	.986
5BI	/	5GP	0.57	0.17	−1.94	.525
5BI	/	5HI	1.56	0.47	1.50	.808
5GH	/	5GP	0.74	0.14	−1.61	.748
5GH	/	5HI	2.05	0.59	2.48	.206
5GP	/	5HI	2.77	0.80	3.54	.010
8BI	/	8GH	0.76	0.23	−0.91	.986
8BI	/	8GP	0.57	0.17	−1.94	.525
8BI	/	8HI	1.56	0.47	1.50	.808

(cont.)

Contrasts			OR	SE	z	p
8GH	/	8GP	0.74	0.14	−1.61	.748
8GH	/	8HI	2.05	0.59	2.48	.206
8GP	/	8HI	2.77	0.80	3.54	.010
5BI	/	8BI	2.17	0.52	3.24	.026
5GH	/	8GH	2.17	0.52	3.24	.026
5GP	/	8GP	2.17	0.52	3.24	.026
5HI	/	8HI	2.17	0.52	3.24	.026

Pronouns

Contrasts			OR	SE	z	p
BI	/	GH	1.95	0.57	2.31	.097
BI	/	GP	1.20	0.35	0.62	.927
BI	/	HI	5.65	1.65	5.92	< .001
GH	/	GP	0.61	0.11	−2.69	.036
GH	/	HI	2.89	0.82	3.74	.001
GP	/	HI	4.72	1.35	5.44	< .001

Imperatives

Contrasts			OR	SE	z	p
BI	/	GH	2.16	0.97	1.72	.312
BI	/	GP	1.11	0.45	0.26	.994
BI	/	HI	10.46	6.99	3.51	.003
GH	/	GP	0.51	0.19	−1.80	.272
GH	/	HI	4.83	3.35	2.27	.105
GP	/	HI	9.42	6.28	3.37	.004

Exclamations

Contrasts			OR	SE	z	p
5BI	/	5GH	1.26	0.58	0.50	1.00
5BI	/	5GP	0.50	0.22	−1.61	.747
5BI	/	5HI	1.52	0.71	0.90	.986
5GH	/	5GP	0.40	0.11	−3.22	.029
5GH	/	5HI	1.21	0.57	0.41	1.00
5GP	/	5HI	3.07	1.37	2.51	.190
8BI	/	8GH	1.26	0.58	0.50	1.00
8BI	/	8GP	0.50	0.22	−1.61	.747
8BI	/	8HI	1.52	0.71	0.90	.986
8GH	/	8GP	0.40	0.11	−3.22	.029
8GH	/	8HI	1.21	0.57	0.41	1.00
8GP	/	8HI	3.07	1.37	2.51	.190
5BI	/	8BI	7.70	3.28	4.80	< .001
5GH	/	8GH	7.70	3.28	4.80	< .001
5GP	/	8GP	7.70	3.28	4.80	< .001
5HI	/	8HI	7.70	3.28	4.80	< .001

You

Contrasts			OR	SE	z	p
BI	/	GH	1.81	0.58	1.87	.242
BI	/	GP	1.18	0.38	0.51	.956
BI	/	HI	5.85	1.91	5.41	< .001
GH	/	GP	0.65	0.12	−2.36	.085
GH	/	HI	3.23	1.03	3.68	.001
GP	/	HI	4.97	1.59	5.03	< .001

We

Contrasts			OR	SE	z	p
BI	/	GH	1.65	0.87	0.95	.776
BI	/	GP	0.88	0.45	−0.25	.994
BI	/	HI	2.78	1.50	1.90	.230
GH	/	GP	0.53	0.15	−2.31	.095
GH	/	HI	1.69	0.91	0.97	.766
GP	/	HI	3.16	1.66	2.20	.124