

Spoken vs. Written or Dialogue vs. Non-Dialogue? Frequency Analysis of Verbs, Nouns and Prepositional Phrases in Bulgarian*

Daniela Assenova

Department of Modern Languages, Uppsala University

Daniela.Assenova@moderna.uu.se

Abstract

In linguistics, the difference between spoken and written language is often interpreted in terms of frequency, meaning the extent of the likelihood that some constructions will occur in written texts, rather than in the spoken form of a language, or vice versa. For Bulgarian, frequency analyses for particular constructions are still rare and the term “frequency” generally remains implicit to the written or the spoken form of the language. In this paper it is argued that the term “frequency” can be meaningful only if it is grounded in an analysis of both written and spoken texts. The primary focus of the study is the frequency of verbs, nouns and prepositional phrases in different types of spoken and written samples. Since neither the written nor the spoken form of Bulgarian, or, indeed, of any language, can be considered homogeneous, it is argued that the differences between spoken and written language can be viewed as differences between dialogue and non-dialogue: i.e., that it is not the mode itself, but rather the structuring of information and its density, which constitutes the difference between the spoken and written modes.

1. Introduction

Most people would agree intuitively that there is a significant difference between spoken and written language. For native speakers, however, the switch between speaking and writing happens automatically, in a given situation, in accordance with the purpose of their communication and social conventions. For example, if two students are communicating with each other during an ongoing class, the final product may result in a written text, in order not to disturb the lecture, but in the structuring of the information in question/answer sequences, their written text would be more similar to an oral conversation, in the sense that the information is structured *ad hoc*, with no

* This paper presents preliminary results of work in progress.

time for editing (since students usually do not want to miss the entire lecture). On the other hand, if a lecturer in the same classroom is reading his or her lecture aloud from a manuscript prepared in advance, which has probably been edited several times and is subject specific, it is more difficult to determine whether the text is spoken or written language. In other words, what are considered spoken and written modes of expression are in essence two forms of communication in the same language, oral and graphic.

The choice of one or another of these two forms of communication usually depends on the situation at hand. What motivates the structuring of the information, i.e., the syntactic parameters, is whether the primary purpose of the communication is interlocutor *interaction* or *information*. In the first case, information is *exchanged* through questions, discussion and argument; in the second, information is *presented*, through the introduction of new material, elaboration, and explanation. The choice of spoken vs. written mode is more situational, while the choice of how to structure the information is more stable. That is, when relating the syntax of written and spoken language, the central issue is *what* is being compared: the spoken vs. written modes as *systems*, the specific *styles* or *registers* within the two modes (academic prose, fiction, journalistic prose, face-to-face conversation, telephone conversation, media speech, etc.), or different *text types* such as *dialogue* and *non-dialogue*? The main focus of this study is whether syntactic constructions in Bulgarian are grounded in a written/spoken opposition, or in a dialogue/non-dialogue opposition. In other words, can dialogue and non-dialogue be “fingerprinted” in syntactic terms as intermediate text types for both spoken and written Bulgarian?

Within the written mode, dialogue in fiction can be distinguished from the dialogue in a play. Both genres represent a simulation of an oral conversation presented through one voice (the writer’s), but they are different in respect to at least one important factor: while the dialogue in a play is written to be performed on a stage, in front of an audience, the dialogue in fiction is most likely to be read by one person at a time. Moreover, the dialogue in a play aims to tell the “whole story”, while the dialogue in a book does not. A similar observation can be made about face-to-face conversation, which can be viewed as dialogue between two (or more) participants, while a discussion in the “media space” is intended for an audience.

The primary objective of a conversation is interaction by which the speech act participants exchange information, while a discussion in front of an audience (visible or not) has a different purpose, in that it is intended to convey information¹. Thus, the purpose of these two kinds of interactions is different, as are their situational characteristics. In fictional dialogue, we can expect a dialogue structure with to some degree equal participation of the involved individuals, and with regular turn-taking.

¹ The use of the opposition interaction/conveyance of information is strongly associated with studies in gender linguistics (e.g. Tannen 1994). In this article all the samples are treated as exchanges or presentation of information, where the gender aspect is not considered.

Such synchronous dialogues are typical of face-to-face and telephone conversations, where some degree of regular interaction between speech act participants is also expected. On the other hand, the dialogue structure of TV talk shows, radio discussions, parliamentary debates or political interviews are much more asynchronous. In these more formal scenarios, the role of the participants is quite different: there is a host or moderator who invites other participants to present information. Thus, although the interaction resembles dialogue in its general structure, the exchange of turns between the involved individuals is not equal. The moderator will usually talk less than the others, while the participants will talk considerably longer, i.e. there will be intervals of non-dialogue. Moreover, the information loading of the texts will be different, since the focus will be to present and elaborate their position.

In syntactic terms, the noun (N) and the verb (V) are the core arguments, the subject and the predicate, in a clause. Their frequency in different texts, whether spoken or written, can indicate how the information is structured. For morphologically analytic languages like Bulgarian, a prepositional phrase (PP) is relevant to both a noun phrase (NP) and a verb phrase (VP), in that it can be either a modifier within the NP or part of the valence of the verb in a VP. The frequency of nouns, verbs, and PPs in written vs. spoken texts is the primary object of this study.

2. Related work

As far as the syntax of spoken and written Bulgarian is concerned, two general positions among Bulgarian scholars can be traced, from before and after the mid-1980's.

The first position, which is based on literary works alone, views the structure of the written dialogue as similar to that of spoken Bulgarian, albeit with more ellipses in the clause and within the sentence (Popov [1961] 1998:62, 112–114).² According to Popov, the most commonly omitted element on the clause level is the predicate, and on the sentence level, the main clause in the answer to a question. Such “grammatically incomplete” (ibid. 112) clauses and sentences are context-dependent and *more frequent* in spoken Bulgarian. In “Gramatika na sǎvremennija bǎlgarski knižoven ezik” (GSBKE 1983:112) the use of “incomplete sentences” is described as frequent in various styles, but most frequent in literature.

The second position on the syntactic structure of spoken and written Bulgarian can be viewed as a change in the focus of the research target, which is defined in Angelova's 1994 monograph as “description of the most characteristic syntactic constructions in spoken Bulgarian” (6). Angelova's phrasing summarizes an important

² In this paper the terms “clause” and “sentence” are used as corresponding to “просто изречение” and “сложно изречение”, respectively, which are the traditional terms in Bulgarian grammar books.

juncture in the research on Bulgarian syntax. In several works (Bajčev & Videnov 1988:219–245; Angelova 1988a, 1988b; Likomanova 1996; Ničeva 1985; Yosifova 1987, 1999; Barakova 2003 etc.), the ultimate goal is to determine the *differences* in the syntactic parameters between spoken and written Bulgarian; however, the emphasis is only on spoken samples. The differences are interpreted in terms of *frequency*, in that the spoken language is spontaneous, contextual and produced in real time. The spoken language is thus characterized as containing more elicitation of verbal and nominal components in mono-predicative syntactic constructions, and showing a preference for asyndetical chaining of syntactic elements in poly-predicative utterances³.

The common denominator of the two approaches is the focus on only one of the two forms of the language, either written *or* spoken, and on the *frequency* of syntactic structures in either written or spoken form. Since the conclusions drawn in those studies are implicit to one of the two forms, this paper examines texts from both forms *in parallel*⁴.

3. Data

The data, presented in Table 1, are divided into two groups: spoken and written samples. Each text sample is limited to approximately 1000 words. The sample size is determined by the desired confidence level and time interval, and by the volume of the text. According to Woods, Fletcher and Hughes (1986:96–111), a confidence interval of 3% is assumed, as well as a confidence level of 95% for a text 10, 000 words in length.

The spoken samples consist of face-to-face and telephone conversations between individuals who know each other well, various types of media speech (TV talk shows, sport commentary, and radio discussions), and public speech (parliamentary debates and political interviews). All the conversations take place in relatively casual situations, such as conversation while doing daily household chores or dinner conversations between friends or relatives, and are informal. The media and public speech texts are recordings from TV and radio broadcasts, political parliamentary debates, and interviews. These are considered formal speech and, more importantly, presentations to or in front of an audience.

³ The change in the focus of the research object also leads to a change in the terminology. For the syntactic units of spoken Bulgarian, “mono-predicative and poly-predicative constructions/utterances” (Bajčev&Videnov 1988:277; Angelova’s 1994:91) are suggested as corresponding terms for the traditional syntactic terms simple and complex sentence.

⁴ Among the research on Bulgarian syntax, studies on *both* spoken and written language are still rare, e.g. see Tiševa & Rå Hauge 2001, 2002; Leafgren 2004, 2007; Fiedler 2008.

<i>Spoken</i>	<i>No. of samples</i>	<i>Written</i>	<i>No. of samples</i>
Conversation		Play	9
Face-to-face conversation	9	Fiction	
Telephone conversation	9	Fiction dialogue	9
		Fiction non-dialog	9
Media talk		Journalistic prose	
TV-talk-show	4	Daily newspapers	4
TV-sport-comment	6	Weekly magazines	6
Radio discussion	3		
Public talk		Academic prose	
Parliamentary debate	2	Articles	7
Political interview	3	Books	5
	<i>Total</i> 36		<i>Total</i> 49
Total spoken and written			85

Table 1. The texts.

The spoken samples were recorded between the years 2005–2008; the written samples were published after 2000. The limitation of the written texts to this period is based on the assumption that it would be inappropriate to compare texts from a different time period with contemporary speech.

4. Spoken vs. written or dialogue vs. non-dialogue?

For purposes of analysis, the texts in Table 1 can be classified according to two parameters: (1) *mode* (written vs. spoken) and (2) *text types* (dialogue vs. non-dialogue). A third parameter can be constructed by conjoining parameters (1) and (2). Parameter (3) allows for the identification of four mode/text types: (3a), written dialogue (WD); (3b), written non-dialogue (WnD); (3c), spoken dialogue (SD); and (3d), spoken non-dialogue (SnD). The relationship among these four types is illustrated in Figure 1.

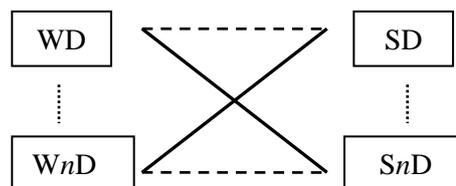


Figure 1. The relation between the four discourse types.

The greatest differences are between WD and *SnD*, and *WnD* and SD, respectively, i.e., where both criteria of mode and text type are different. On the other hand, if only one parameter is varied, different junctures can be identified within the same mode or within the same text type. The dotted line between WD and *WnD* and between SD and *SnD* is “readable” as a continuum within a single mode, e.g., registers; the dash line between WD and SD and between *WnD* and *SnD* represents correspondence in text type, but difference in mode. The issue is whether the dotted line (registers) or the dash line (correspondence in text type) has more influence on the type of syntactic construction used.

A spoken and a written text presented in the same mode can be difficult to distinguish:

(1)⁵

A1: Може ли да ***се говори*** за затваряне на тази ножица?

B1: Напротив. Икономическото развитие след двехиляди и първа година ***разтвори*** тази ножица между много бедни и много богати. В този смисъл това развитие, опиращо се на спекулативна основа, ***даде*** път да ***се развива*** черния и сивия бизнес в България, да ***се изперат*** пари, да ***се легализират*** бизнеси. Общо взето либералното управление ***беше*** услужливо към тези бизнеси и цялата българска икономика постепенно ***получи*** такива двигатели за своето развитие. От 2006 г. насам, това нездравно развитие на практика непрекъснато ***показваше***, че балоните, които ***е надуло*** ще се спукат. И ето я българската фондова борса. Най-добрият показател за балони. От две хиляди на триста, основния ѝ индекс.

A1: (‘Is it possible to talk about a closing of these scissors?’)

B1: (‘On the contrary. The economic development after year 2001 opened the gap between the very poor and the very rich. In this sense, this development, grounded on a speculative basis, opened the way for development of the black and gray businesses in Bulgaria, for money-laundering, for legalization of such

(2)

”Според кмета и неговата администрация депото не ***отговаря*** на нормите за безопасност. Проверка РДНСК – Стара Загора, от юли т.г. обаче ***установи***, че депото ***е*** изпълнено според изискванията на проекта и не ***е*** незаконно. Личният двубой на кмета с “Енел” ***достигна*** своята кулминация миналата година, когато жители на село Медникарово в присъствието на Тонев и председателя на общинския съвет в Гълъбово ***спретнаха*** митинг, на който демонстративно ***изгориха*** сламени кукли с образите на двамата шефове на “Енел” – Майк Фостър и Енрико Виале. Кадрите с горящите чучела ***напомнят*** за наказателните акции на ку-клукс-клан срещу чернокожите в Америка през XIX век.”

(‘According to the mayor and his administration the landfill does not meet the safety standards. The inspection by RDNSK - Stara Zagora, in July this year however, found that the landfill is implemented according to the project requirements and is not illegal. The personal battle between the mayor and “Enel” reached its peak last year when the residents of the village Mednikarovo in the presence of Tonev, the chairman of the municipal council in Gălăbovo made a public demonstration, where they

⁵ The examples are given in a suitable form for the analysis, normalized according to the written norm. Instead of a dash (“–”) at the beginning of a line of dialogue, the participants are marked with A, B, C etc. A number is associated with each turn of each participant, e.g. A1, B1, C1, B2, etc. The texts are given in Bulgarian and translated into English. In order to be more “readable” the verbs in Bulgarian are marked in bold and italics.

businesses. In other words the liberal government was very suitable for those types of businesses and the entire Bulgarian economy gradually received such engines for its development. Since 2006, this unhealthy development showed that the balloons, which were blown up, are going to burst. And here we have the Bulgarian Stock Exchange. The best indicator for the balloons. The main index went down from 2,000 to 300.’)

demonstratively burned straw dolls with images of the two heads of “Enel” - Mike Foster and Enrico Viale. The scenes with the burning dummies remind of the criminal actions of the Ku-Klux-Klan against the black people in America during the 19th-century.’)

Example (1) is a sample from a radio discussion; (2) is an excerpt from a newspaper article. Now consider the following parallel samples of oral and written language:

(3)

C1: Кое *е* това преди? **Кажу** сега!
‘What do you mean before? Tell me now!’

E1: Онуй преди войната.
‘That before the war.’

C2: Кое преди бе, Дони?
‘What before the war, Doni?’

D1: Това, което *е* преди. До това в Сараево.
‘This that was before. Until that in Sarajevo.’

C3: Не, щото преди *е* много общо.
‘No, because “before” is too general.’

E2: **Знаеш** ли го кой *е* Гаврило Принцип?
‘Do you know who Gavrilo Princip is?’

D2: Да, той **уби** Франц Фердинанд.
‘Yes, he murdered Franz Ferdinand.’

C4: И ти **беше** там?
‘And you where there?’

D3: Бъхъ. **Няма да** го *пиша* на български.
‘Aha. I won’t write it in Bulgarian.’

E3: К’во?
‘What?’

C5: Не, обаче **искам** като *говориш* да *говориш* като нормален човек. Преизказно наклонение и други. **Чу** ли?
‘No, but I want you to speak as a normal person. Narrative mode and so on. Did you get it?’

E4: А кой **уби**?
‘And who murdered?’

D4: Франц Фердинанд. **Казах**.
‘Franz Ferdinand. I just said.’

(4)

F1: А щом *е* с кола, значи **трябва** да **бунне** някого.

‘Well, if he is with a car, it means he has to bump into somebody.’

G1: Нея?
‘Her?’

F2: Окей. Тя каква *е*?
‘Okay. What is she like?’

G2: Не **пие**, не **пуши**, на 35?
‘Doesn’t drink, doesn’t smoke, 35 years old?’

F3: Дали да не *е* малко по-така?
‘Shouldn’t she be a little bit different?’

G3: А ти какво **искаш**: **пие**, **пуши** и *е* на пет? Обикновена ни **трябва**.
‘And what do you want: she drinks, smokes and is 5 years old? We need an ordinary one.’

F4: **Може** да *е* така: **пие**, **пуши** и всеки **иска** да **знае** на колко години *е*, а?
‘It could be like this: she drinks, smokes and everybody wants to know how old she is, right?’

G4: **Може. Става**.
‘It could be. It works.’

F5: И **бута** я...
‘And he bumps into her ...’

G5: Ама лекичко.
‘But just a little.’

F6: Тя *е* виновна.
‘She is guilty.’

G6: Защото **пресича** неправилно.

‘Because she crosses improperly.’

At first reading, the above two examples are virtually indistinguishable⁶. Example (3) is from a face-to-face conversation between two parents (C, the mother, E, the father) and their son (D); (4) is from a written dialogue in a short story between two male characters, F and G. If examples (1) and (2) are compared with (3) and (4), it is clear at first glance that there is a significant difference in their syntactic structure. While the first two examples give a more linear impression by structuring the information in longer syntactic units, the information in examples (3) and (4) is processed in smaller units, i.e. “step by step”. A similar first impression is also obtained by a parallel reading of samples (1) and (3), both in spoken mode, or (2) and (4), both in written mode.

Sample (3) contains seven questions (C1, C2, C4, C5, E2, E3, E4), whereas sample (1) has only one (A1). All three participants C, D and E in (3) contribute almost equally to the conversation, with C having five turns, and D and E having four turns apiece. In (4), the two participants F and G are involved equally, and have taken 6 turns each. Both texts are interactive in the sense that they contain questions relating to the immediate context. In samples (3) and (4), the information is realized through either *k*-questions⁷ or questions structured by intonation, repeating a word from the previous turn (G3, F4). In sample (1), there are also two participants, A and B, but their contributions are uneven: A asks the only question and in this way initiates the presentation, while B’s turn can be compared to the expository journalistic (written) text in (2), where no questions occur. Thus, for information structuring purposes in both dialogue and non-dialogue, the frequency and type of questions are central: in dialogue the information is processed in small portions through questions, while in non-dialogue it is given at once and further elaborated.

The information density in the four texts can also be grouped according to dialogue and non-dialogue. In (1) and (2), the distribution of lexical categories is as follows: verbs 9% and 6%, and nouns 23% and 36%, respectively. The percentages in (3) and (4) are verbs 19% and 33%, and nouns 11% and 2%, respectively. The dialogue/non-dialogue distinction with respect to verbs is found not only in the frequency of verbs,

⁶ An experiment using all 4 examples was made in 2010. The four texts were presented to 142 subjects of high school to Ph.D. educational levels and from ages 16 to 74. The subjects received all four samples without punctuation or capital letters. Their task was to determinate the mode and type of each text, and then to add all necessary punctuation marks and capital letters. One preliminary result from the experiment is that the subjects found the relationship between written and spoken difficult to establish in (3) and (4). Sixty-nine percent of the subjects incorrectly considered the written example (4) to be spontaneous spoken, whereas 94% marked both (3) and (4) correctly as dialogue. Samples (1) and (2) were also correctly marked as spoken non-dialogue and written non-dialogue by 87% of the subjects. Since the experiment is still in progress, no general conclusions can be drawn at this time, although some specific tendencies are already noticeable.

⁷ In Bulgarian, several interrogative pronouns begin with the letter <к>, e.g. *кой, коя, кое, кои, какъв, каква, какво, какви, къде* etc. The *к*-words correspond to the well known term *wh*-words in English. The term *к*-questions is used to make a distinction from *yes/no*-questions, which are constructed with the interrogative particle *ли*.

but also in verb types. In (1) and (2), most of the verbs are transitive (*разтвори, даде, получи* etc, *отговаря, установи, достигна* etc) and in reflexive form (*се развива, се изперат, се легализират*) while in (3) and (4) intransitive and modal verbs are preferred. Thus, in the giving of information in non-dialogue, syntactic structures with two overt referents subject and object, are needed, while for the interactive text types, in dialogue, one referent is often enough. Moreover, when the referent is one of the speech act participants, it is incorporated in the personal ending of the verb (notice the frequency of verb forms for first and second person in examples (3) and (4)). When the referent is not one of the speech act participants, structures with the copula verb *съм* are preferred (compare the frequency of the verb *съм* in (3) and (4) with its occurrence in (1) and (2)).

The differences between (1) and (2), and between (3) and (4), are also expressed through the use of PPs. In the non-dialogue texts, the number of PPs is higher, and they are used preferably as modifiers in the NPs (*икономическото развитие след две хиляди и първа година, тази ножица между много бедни и много богати, личният двубой на кмета с "Енел", сламени кукли с образите на двамата шефове на "Енел" – Майк Фостър и Енрико Виале*), while in the dialogue texts, the number of PPs is lower, and they are either “free” (*До това в Сараево*), or related to the V by giving additional information as adverbial (*да говориш като нормален човек, е с кола*).

Thus dialogue and non-dialogue can be distinguished in respect to the frequency of V, N and PPs. According to Biber (1988:13) the structuring of information leads to the “co-occurrence” of linguistic features and their frequency is grounded in the functional characteristics of different types of texts, described as dimensions of variations among spoken and written English. Moreover, in Biber (2004:21) the author presents results that conversations tend to be “either ‘informational’ or ‘interactive’ but not both”. In his study this is referred as informational-focused vs. interactive discourse.

Figure 2 below shows the frequency of Vs, Ns and PPs when the focus is on mode (i.e., written vs. spoken language), Figure 3 shows the results for text type (i.e., non-dialogue vs. dialogue), and Figure 4 shows the results when the texts are distinguished according to both mode and type. The numerals in the figures represent average values for the different categories: for example, the result labeled “play” gives the average value for all nine plays that were included, while the result for “TV-talk-show” gives the average value for all 4 talk-shows, etc.

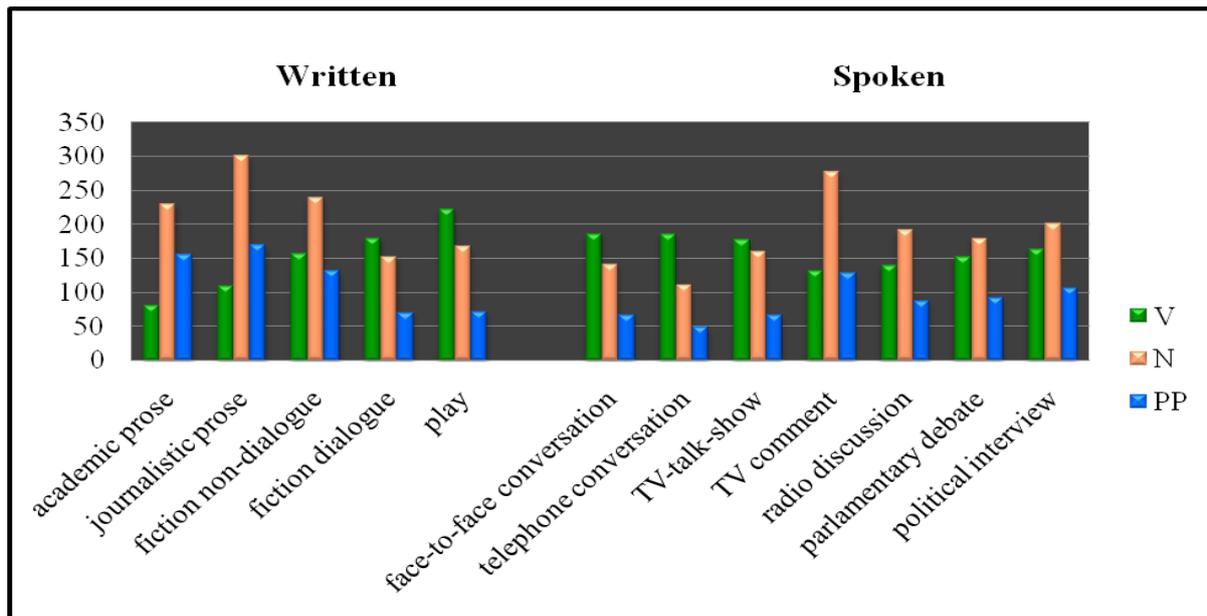


Figure 2. The frequency of V, N and PP in spoken and written texts.

Figure 2 displays the differences in the frequency of V, N and PP, for the different categories, but it is difficult to make a general conclusion when the only distinction is that of mode. Fictional dialogues and plays are different from the rest of the written material in that they contain more Vs than Ns and considerably fewer PPs than the rest of the written texts; similarly, TV commentaries, radio discussions, parliamentary debates and political interviews differ from the rest of the spoken material, containing more Ns than Vs.

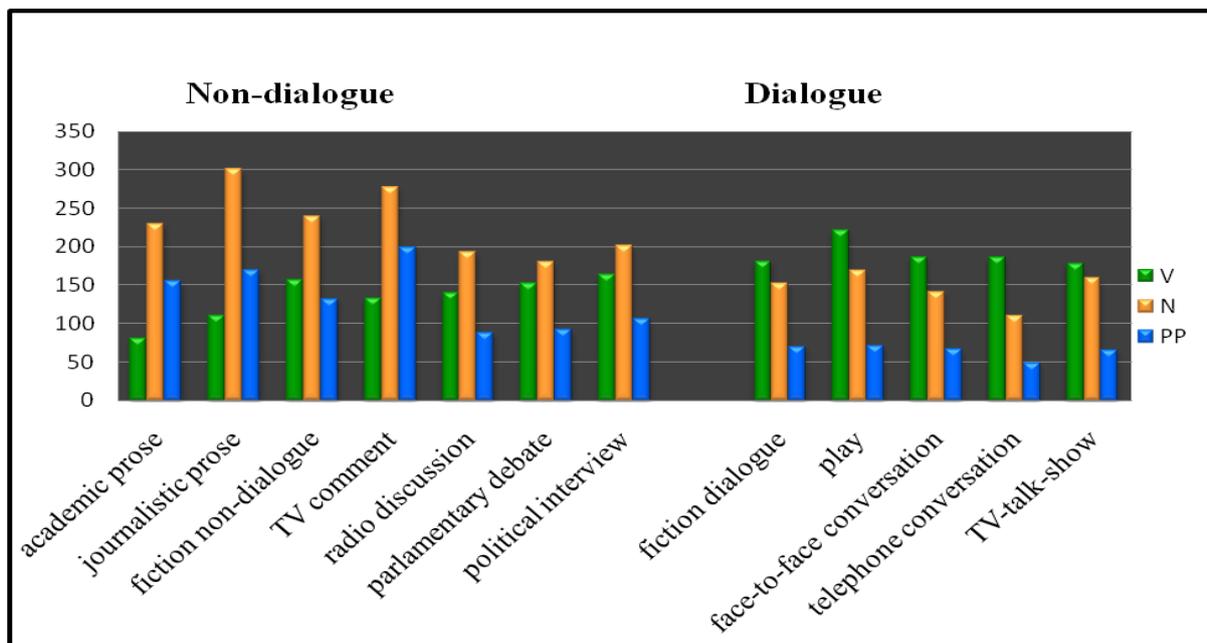


Figure 3. The frequency of V, N and PP in dialogue and non-dialogue.

When the distinction is by text type, as in Figure 3, irrespective of written or spoken form, there is a strong correlation between the frequency of verbs vs. nouns. In non-dialogue, there are considerably more Ns than Vs, whereas in dialogue the relation is reversed. A similar pattern is displayed in the frequency of PPs, which are considerably less frequent in dialogue. Since PPs can be related to both nominal and verbal elements, further analysis is necessary in order to determine to what extent the frequency of PPs in dialogue is related to the frequency of Ns⁸.

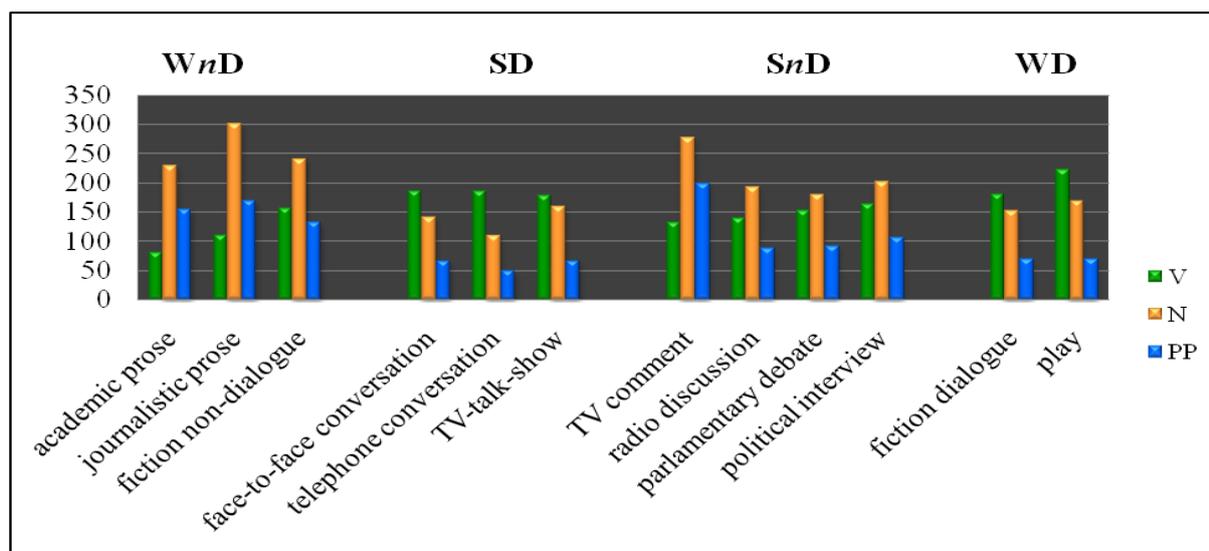


Figure 4. The frequency of V, N and PP with respect to mode and type.

Figure 4 shows that the greatest differences lie between written dialogue (WD) and spoken non-dialogue (SnD), and between written non-dialogue (WnD) and spoken dialogue (SD), respectively, i.e., when both mode and text type are different.

5. Discussion

So far the following four points can be made: (i) neither spoken nor written Bulgarian is homogeneous in respect to the frequency of Ns, Vs and PPs; (ii) the distinction between dialogue and non-dialogue is significant for both forms of the language, because (iii) there is a correlation between the frequency of nominal vs. verbal elements and the dialogue vs. non-dialogue opposition, but not the spoken vs. written opposition; and, based on these observations, (iv) in order to be more meaningful,

⁸ One complication, for example, is the classification of PPs with the preposition “на”, which in Bulgarian can function as (i) a possessive marker (as in Това е къщата на Иван/This is Ivan’s house), (ii) a dative marker (Давам книгата на Иван/I am giving the book to Ivan) and (iii) a preposition in an adverbial (Иван стои на улицата/Ivan is standing on the street). In (i) the PP is a post-modifier in a NP, while in (ii) and (iii) it is an object and an adverbial, respectively (i.e. an argument and an extension of the VP). It must be also considered that in respect to (ii), the preposition “на” is dropped when replacing the N with a short dative pronoun in Bulgarian. Thus, relying on the quantitative analysis alone is not enough.

frequency studies of lexical categories should take into account whether the text samples under study are dialogue or non-dialogue samples.

Moreover, the relation between the spoken and the written form of a language does not arise only from the situational characteristics of the texts, that is, the fact that spoken language is more contextual, spontaneous, and in real time. With respect to the situational characteristics, a striking example is the result for TV sport commentaries (see Figure 4). The mere fact that these commentaries are delivered in an oral form is not enough. The context for both the TV-audience and the commentators is identical, in that both are watching the same game, the video is delivered in real time, and neither the TV audience nor the commentators is prepared in advance. What makes the results for sport commentary very similar to those for journalistic prose and different from the results for all other oral texts is that the purpose of the sport commentary is to give additional information about the events taking place, which puts the focus primarily on the information. This again illustrates that the information structuring in Bulgarian is more dependent on the text-type opposition dialogue/non-dialogue, than on the opposition between written and spoken form.

References

Angelova 1988a: Ангелова (=Ликоманова), И. "Някои наблюдения върху едно явление от устната реч". *Славистичен сборник*. София: БАН. www.slav.uni-sofia.bg/bgspeech/sites/en/publ/bgsp_il_niakoi.pdf, 2010-11-20.

———. 1988b. "Организация на репликата в съвременната българска разговорна реч". *Български език* 6, 525–529.

———. 1994. *Синтаксис на българската разговорна реч (в съпоставка с руски, чешки, полски език)*. София: Университетско издателство.

Вайчев & Виденов 1988: Байчев, Б., Виденов, М. *Социологично проучване на град Велико Търново*. София: Наука и изкуство.

Barakova 2003: Баракова, П. "Наблюдения върху полипредложните конструкции". *Српски језик* 8, www.slav.uni-sofia.bg/bgspeech/sites/publ/Polipredlozni%20konstrukciji.pdf, 2010-11-20.

Biber, D. 1988. *Variation across speech and writing*. London: Cambridge University Press.

———. 2004. "Conversation text types: A multi-dimensional analysis". *JADT:04.7^{es} Journées internationales d'Analyse statistique des Données Textuelles*, 15–34.

Fiedler, G. E. 2008. Bulgarian adversative connectives. Laury, R. (ed.), *Crosslinguistic Studies of Clause Combining: The Multifunctionality of Conjunctions*. Amsterdam: John Benjamins, 79–98.

GSBKE 1983: ГСБКЕ. Попов, К. (гл. ред.) *Граматика на съвременния български книжовен език*, т. 3 (*Синтаксис*). София: БАН.

Josifova 1987: Йосифова, Р. ”Словоред на съгласуваното определение в книжовно-разговорната реч”. *Език и литература* 2, 69–77.

———. 1999. ”По въпроса за синтактичните особености на книжовно-разговорната реч”. *Български език* 4–5, 76–81.

Leafgren, J. 2004. “Bulgarian Relative Clauses and Communicative Complexity”. *The Slavic and East European Journal* 48–2, 275–291.

———. 2007. “The Use of Verbal Adverbs and other Verbal Adverbials in Formal versus Informal and Written versus Spoken Bulgarian”. *Съпоставително езикознание* 1, 25–43.

Likomanova 1996: Ликоманова (=Ангелова), И. Многократността на изразните средства като особеност на българската разговорна реч. *Проблеми на българската разговорна реч*, кн. 3. В. Търново, 116–121.

Ničeva 1985: Ничева, К. ”Наблюдения върху разговорната реч на българския език”. *Български език* 3, 201–210.

Rorov 1961: Попов, К. *Съвременен български език. Синтаксис*. София: Наука и изкуство.

———. 1998. *Синтаксис на съвременния български език*. В. Търново: Абагар.

Tiševa & Rå Nauge 2001: Тишева, Й., Ро Хауге, Х. Съюзи за противопоставяне във функцията на прагматични частици. *Проблеми на българската разговорна реч*. кн. 5. В. Търново, 242–252.

———. 2002. Заемане на прагматични частици. *Проблеми на социолингвистиката*, т. 7. *Билингвизъм и диглосия – съвременни проблеми*. София: ”Делфи”, 10–17.

Woods A., Fletcher P. & Hughes A. 1986. *Statistics in Language Studies*. London: Cambridge University Press.