UDC 81’42

A. D. Mikulinskiy
HSE University

FUNCTIONAL MODELS OF ELEMENTARY DISCURSIVE UNITS IN RUSSIAN ESPORTS COMMENTARY

The paper is devoted to the issue of the local structure modeling of the eSports commentary spoken genre on an example of the Dota 2 computer discipline. ESports commentary is a spontaneous and creative speech aimed at describing of what is happening on the computer-gaming field. The main factors that force us to study it are the high popularity, the influencing nature of speech, as well as the lack of scientific attention. A theoretical and methodological framework of the study contains the elements of the structural and cognitive-discursive approaches in linguistics. The key research methods are language modeling, analysis of the local discourse structure in the cognitive perspective and quantitative analysis with regard to the corpus-based approach. The statistics were calculated for the sample with the total volume of 41 minutes and 30 seconds. The speeches belong to 14 Russian eSports commentators; they were delivered in 2017–2019. As a result, we have obtained a set of relevant patterns that represent characteristics of typical elementary discursive units (syntagmas) in the Dota 2 eSports commentary. These include a quick pronouncing, boundary pauses absence, frequent nuclear accent presence and embodiment of nouns and verbs. One part of the statistics represents general features of the Russian spoken discourse and human consciousness (e. g. the frequent use of absolute pauses); the other part correlate with a specificity of the situation in which the commentary is produced (e. g. the frequent use of short structures).

Keywords: eSports, commentary, Dota 2, language model, functional model, elementary discursive unit, spoken discourse.

А. Д. Микулинский

ФУНКЦИОНАЛЬНЫЕ МОДЕЛИ ЭЛЕМЕНТАРНЫХ ДИСКУРСИВНЫХ ЕДИНИЦ В РУССКОЯЗЫЧНОМ КИБЕРСПОРТИВНОМ КОММЕНТАРИИ

Статья посвящена проблеме моделирования локальной структуры киберспортивного комментария на примере выступлений ведущих в сфере компьютерной дисциплины Дота 2. Киберспортивный комментарий представляет собой спонтанную и творческую речь, направленную на описание событий, происходящих на компьютерно-игровом поле. Основными факторами, вынуждающими нас обратиться к изучению данного жанра, являются его высокая популярность и степень воздействия

на зрителя, а также малая изученность с лингвистической точки зрения. Теоретико-методологическая рамка исследования содержит элементы структурного и когнитивно-дискурсивного подходов в лингвистике. Основными методами являются языковое моделирование, анализ локальной структуры дискурса в когнитивной перспективе и количественный анализ в рамках корпусного подхода. Статистики были подсчитаны для выборки объемом 41 минута и 30 секунд; проанализированная речь принадлежит 14 киберспортивным ведущим, выступавшим в 2017–2019 годах. В результате анализа была получена совокупность паттернов, которые позволяют судить о характеристиках типичных элементарных дискурсивных единиц (синтагм) в киберспортивном комментарии. Эти характеристики включают: быстрое произнесение, отсутствие внешних пауз, частое появление несущих акцентов, существительных и глаголов. Одна часть статистик объясняется «нормой» русского устного дискурса и особенностями человеческого сознания (напр. частое использование абсолютных пауз в речи). Другая часть указывает на специфику сферы и ситуации, в которой производится рассматриваемый дискурс (напр. частое использование коротких структур).

Ключевые слова: киберспорт, комментарий, Дота 2, языковая модель, функциональная модель, элементарная дискурсивная единица, устный дискурс.

Introduction

The virtual environment and eSports have long been a part of many people's lives. Having begun to flourish more than twenty years ago, the field of eSports, and online eSports in particular, has got a rapid and serious development at the beginning of the XIX century becoming "…a fundamental element in today's digital youth culture" [Wagner, 2006, p. 437; Pankina, 2016, p. 34]. Within the eSports discourse, genres similar to those established in the sports field years ago, such as interview, news, official announcement, analytics and commentary, have come into being.

The profession of an *eSports commentator* (*caster*) is one of the most crucial types of activity within the eSports community. Casters' participation at the tournament broadcasts primarily contribute to attracting the viewers and keeping them at the screen. As Randhawa (2015) truly notices, "all together, professional gamers, audience members, and commentators present a dynamic understanding to video games as a performative medium" [p. 16]. In this plain triad, commentators perform a mediation role: they deliver information emerging from gamers' actions on the computer-playing field to the viewer. Although the informing function is the most significant, we should acknowledge that commenting is something more than just dry game facts enumeration. Indeed, we should not forget

about such tasks as audience entertainment, self-presentation and image development [Sukhodimtseva & Sukhodimtsev, 2016; Li et al., 2020]. In this way, the aspect under the focus in the current paper is commentators' language as a tool for delivering game facts, for evoking emotions in the viewer and bringing fame for the community members.

Notably, what makes the eSports commentary ***relevant*** for studying is not only its special functionality, but the discursive space and the conditions in which it is produced. Combining sports and gaming terminology, the features of monologue and dialogue speech, *eSports commentary* represents a swiftly-generated, unprepared and creative spoken discourse constrained by media-community norms [Mikulinskiy, 2020, p. 95–96; Himik, p. 461]. These factors both expand the range of scientific studies to be conducted on the casters' speech material and increase the applicability of results. In this research, however, we limit ourselves to answering ***the following question***: when Russian eSports commentators strive to stay within the genre, to fulfill their professional duties and not to fall short under the influence of conditions, what specific discursive structures do they primarily rely on? By saying "specific discursive structures" in this context, we are referring to quite concrete phenomena: elementary discursive units and their constituents.

Elementary discursive units (hereinafter, *EDUs*), which in some linguistic theories are also referred to as "syntagmas", are elementary building blocks of the genre. Given the fact that any discourse can be regarded from the global and local perspectives, our focus falls on the *local (inner) structure* of the eSports commentary, the elements of which are EDUs in different forms and relations [Carlson et al., 2003, p. 86; Kibrik, 2003, p. 35]. Since EDUs are very diverse in their organization and semantics, the only adequate way to represent and study them seems to be through *modeling*. Thus, ***the aim of the study*** can be formulated as follows: using the functional model of the local structure of the Russian eSports commentary, to identify the EDUs characterized by the greatest relevance, and make assumptions about their functional load in the genre under consideration.

The stated aim implies performing ***the following objectives***: 1) to define the philosophy of the study and outline its theoretical and methodological framework; 2) to select, collect, transcribe and annotate the material of the study, the oral speech of Dota 2 eSports commentators; 3) to create and describe related aspects of models of the eSports commentary local discourse structure; 4) to perform a quantitative analysis of the EDU *patterns* (submodels) in the material; 5) for the most frequent patterns, make assumptions about the role that they play in the current genre.

Dota 2 has been chosen as a ***computer discipline*** on the example of which the current research is conducted, and the reason for this choice can be justified. First consideration is the high online and popularity of the game, the data on which can be found on the Steam website (https://steamcharts.com/app/570). Secondly, the Dota 2 professional and fan communities, including Russian commentators› circle, are highly developed, whereas access to official tournaments and various international competitions is free. This simplifies the procedures for material search and selection.

Methods and material

The material of the study is *41 minutes and 30.17 seconds* of oral speech that belongs to 14 Russian Dota 2 commentators who performed during 9 eSports official tournaments conducted from 2016 to 2019. Such a period was chosen due to the fact that the game has undergone significant changes before 2016 and after 2019, and it could affect the commentary structure. The audio recordings were derived from the competitions broadcasts available at YouTube.com website. The material, thus, contains *14962 tokens* distributed by *1819 EDUs*.

It should be noted that the described sample is just a step on a way to the broader goal: the creation of the *Russian eSports commentary spoken corpus* (*RECSC*). The term "corpus" is designed to emphasize that the sample is not a simple collection of the graphically registered audio recordings. Its compilation is subject to the principles of representativeness, balance, homogeneity, electronic representation and presence of annotation (including the designations for transcription). And since a number of these principles are not observed for the material of the current study, the existing sample cannot be called a corpus so far. The table 1 below contains the information on the comparison of the research material ("observed") with a future corpus sample ("required"), on the basis of which more reliable conclusions could be drawn.

As you can see from the table, the relative representativeness, balance and homogeneity of the sample come speech selection with regard to three key criteria: number of interlocutors (C1; single vs. pair performances), speech belonging (C2; the commentators) and game event type (C3; active vs. passive events). For the C1, the shares of 25 % and 75 % were established based on the analysis of 100 random video recordings of the Ruhub YouTube channel for 2018: about three quarters of the recordings turned out to be paired performances. The establishment of equal shares for the individual casters (C2) is connected with the assumption that syntactic

Table 1. **Data on the current research sample and its required state**

Principle	Criteria/Type	Description
Volume sufficiency	–	**Required**: ~100 min. of speech, ~25.000 tokens; ~4000 EDUs **Observed**: ~41 min. of speech; 14962 tokens; 1819 EDUs
Representativeness, balance & homogeneity	**Interlocutors number (C1)**	**Required**: *≈25 % on single and ≈75 % on pair performances* **Observed**: 21.4 % vs. 78.6 %
	Speech belonging (C2)	**Required**: equal shares for each caster in the sample **Observed**: for 11 casters, the shares are approximately equal, with maximum deviation of 30 seconds. For 3 casters, there is a strong deviation of 3 minutes or more.
	Game event type (C3)	**Required**: equal shares for active and passive events **Observed**: 48.4 % vs. 51.6 %
Representation	–	Wrapped in XML tags, access is carried out through a Python program
Presence of annotation	**Metadata**	Displays data on corpus, on its subsections, on tournaments and on audio segments
	Transcriptional annotation	Displays 25 discursive phenomena
	Morphosemantic annotation	Displays verbals, tokens, lemmas, part-of-speeches, multi token words, part-of-speeches and lemmas of multi token words.

and semantic characteristics of commentary seem to greatly vary from one person to another. And finally, various game situations (C3) also have a serious impact on the discourse production. The game recordings, due to their great length, cannot be entirely put into the sample; however, the in-game events variety should certainly be taken into account. Thus, aiming to classify game events, we have divided them into *active* ones (chases, battles and attacks) and *passive* ones (all other points, for example, gold farming). Due to the impossibility of adequate calculation, their shares were also set equal.

The sample is also accompanied by the annotation, which acts as the most general functional model within the study. It includes three types: metadata, transcriptional data and morphosemantic annotation. The last two act as submodels. For now it is enough to specify that a number of phenomena are displayed in each type, and some of them form the EDU patterns, which are the subject of our attention.

A ***theoretical and methodological framework*** of the study contains the elements of the *structural* and *cognitive-discursive* approaches in linguistics. Using the term "(functional) language model", we primarily refer to the works by Apresjan (1966), Baranov (2001) and Kravtsova (2014), whose ideas partly date back to the classical descriptivists and functionalists' proceedings. Our view on the notions of discourse, EDUs and discursive transcription is fully consistent with Kibrik's (2003) understanding that is based, in particular, on proceedings by Chafe (1994), Levelt (1993), Du Bois (1992), Du Bois et al. (1993) and Carlson, Marcu and Okurowski (2003). The annotation was approached through the lens of studies by McEnery and Hardie (2012), Garside, Leech and McEnery (2013) and Kuebler and Heike (2018).

Selection of the material consisted in *stratified random sampling*. With respect to the C1 and C2 (see the table 1 above), we created a list of casters whose speech needs to be sampled. Then we randomly chose the video records that contain their single and pair performances from YouTube.com. And finally, aiming to meet the C3, we extracted audio segments with commentators' speech during active and passive events with respect to the volume. When extracting the recordings, "Sound Forge 9.0" program was applied for their adequate processing.

Working with the material, namely its transcription, annotation and patterns counting, was performed with a variety of methods and methodics. The main place among them is occupied by an *analysis of the local discourse structure in the cognitive perspective* developed and described by Kibrik, Podlesskaya and their colleagues (2009). This method, being complex, assumes the application of other methodics such as expert assessment of pitch frequency and tone directions, determining the pauses duration and types, linguistic units semantic load analysis, introspection and many others, including general scientific methods, e. g. comparison and classification. The result of its application was discretely reflected speech in the form of EDU sequences. Further, the transcribed material was wrapped into the XML annotation by means of the Python programming language. This included *automatic part-of-speech tagging* with subsequent *manual processing* (correction of the errors in tags assignment via the *dictionary* method). Lastly, following the *corpus-based* study philosophy, a *calcula-*

tion of relative frequencies and *n-gram analysis* were performed to determine the degree of relevance of the EDU patterns. When working with the material, we also used a number of programs and electronic resources: Sound Forge 9.0, Notepad++, Speech Analyzer 3.0.1 and Google colab.

Literature Review

The theoretical and methodological framework of the research contains four general aspects: eSports studies; language modeling; discourse analysis and discursive transcription; and sample annotation. Let us consider the prerequisites, notions and methods of each aspect, around which the study is built and which determine the gap formed.

1) **ESports: issue field & the gap**. While in the foreign scientific literature scientific interest in eSports began to arise in the first decade of the 20th century [cf. Wagner, 2006], in Russia, apparently, humanities became interested in this phenomenon much later. Following Boguslavskaya et al. (2018), we understand *eSports* as a competitive video gaming in the virtual environment [p. 104]. As games vary, eSports studies are typically bound to specific team disciplines (e. g. Dota 2). Each discipline has its own *media community* responsible for the event coverage, and eSports commentators are its integral part [Boguslavskaya et al., 2018; Sukhodimtsev & Sukhodimtseva, 2016]. Both foreign and Russian literature touch on the topic of eSports casters' activity, cf. work by Kempe-Cook, Sher and Su (2019) that describes casters' practices; study by Sukhodimtsev (2016) where the relevance of eSports among the youth seems to be confirmed; work by Hamari and Sjöblom (2017) that investigates motivations of eSports watching, etc. Linguistic and discursive aspects of the eSports commentary, however, seem to be rarely touched [cf. Zaripov, 2016]. What is more, apparently, no scholar addressed eSports commentary from the cognitive-discursive perspective and researched its production and local level structure. Given the relevance of such a study outlined in the introduction, we have just stated about its ***scientific novelty***.

2) **Language modeling**. Modeling is a method that consists in replacing the original with a model in order to study the former through the latter. Yartseva et al. (1998) and Apresjan (1966) define *language model* as an artificially constructed device, tangible or intangible, which purpose is to simplistically imitate the behaviour of language, as it is not directly observable. For any language model, the following is true: 1) it is an ideal object that transmits *the most significant* (for the researcher) *structural* and *functional* properties of language; 2) the language model in itself, therefore, is *formal* and *functional*; 3) it has *predicative* and *explanatory*

nature; 4) language models can be used for studying *discourse production* [Apresjan, 1996, p. 79–80; Baranov, 2001, p. 6–7; Kravtsova, 2014; Kibrik & Podlesskaya, 2007, p. 31].

The proceedings by Apresjan (1966) and Yartseva et al. (1998) represent the classification of language models that seems to be generally acceptable. However, we are interested in a specific type: *a model of analysis*. This type of language model falls into the category of *functional models* that are able to correlate the form and the content of linguistic units; it represents a "finite number of rules for an infinite number of sentences analysis" [Apresjan, 1966, p. 106; our translation, A. D.]. According to Revzin (1977), such a model is designed to output syntactic and morphological categories that the constituents belong to. Earlier Apresyan (1966) suggested dividing language models into *syntactic* and *semantic* based on kinds of output records: structural vs. semantic.

3) **Discourse analysis and discursive transcription**. *Discourse analysis* as an autonomous branch of linguistics was established at the end of the XX century. Currently, there is a large variety of approaches to discourse investigation. Following Kibrik (2003), the paper focuses on the *functional-cognitive approach* to discourse understanding where the speaker's perspective (speech production) is taken as a basis. Let us briefly list the main ideas that we adhere to.

In Kibrik (2003), *discourse* is understood as the unity of communication as a procedure of knowledge exchange and text produced in the process of it [p. 4]. Van Dijk and Kintsch (1983) notice that discourse as a complex phenomenon has two sides or forms: global and local. At the local level that interests us most, discourse is represented in the form of sequentially generated quanta or elementary discursive units [Kibrik 2003, p. 35]. With regard to Chafe's (1994) cognitive view, EDU reflects the focus of speaker's consciousness and carries rheme. From the position of Austin's Speech Act theory (1963), sequences of EDUs form speech acts (*sentences*). Their flowing is needed for communicative intentions realization [Kibrik & Podlesskaya, 2009, p. 95].

According to Du Bois (1992), discursive transcription is a procedure that implies theoretical comprehension of a discourse flow including the selection of phenomena for displaying. Apart of prosodic characteristics of speech, discursive transcription reflects other language aspects such as grammar and semantics, as well as situational, extralinguistic information [p. 73–74]. Thirteen years ago, Kibrik and his colleagues developed a cognitive-*discursive transcription* system, a set of related rules and principles for discrete reflection of Russian spoken discourse [namely, Kibrik & Podlesskaya, 2009]. The groundwork of this system is particularly repre-

sented by the proceedings by Austin (1963), Van Dijk and Kintsch (1983), Mann and Thompson (1988), Chafe (1994), Du Bois (1992), Du Bois et al. (1993), Levelt (1993), Tomlin et al. (1997) and Kibrik (2003). Their system is designed to reconstruct speaker's cognitive processes and operates with a number of special notations. These include "*accent*" (prosodic emphasis), "*nuclear accent*" (stressing the syllable of a word that carries information important at the current step of production); "*accent tone direction*" (tone behavior within accents), "*pause*" (a period of silence (*absolute p.*) or hesitation (*filled p.*) at the border of an EDU (*outer p.*) or inside it (*inner p.*)); "*illocution*" (the communicative purpose) and many other [p. 65-94]. In a sounding speech flow, EDUs are distinguished with accordance to a set of criteria. Kibrik and Podlesskaya (2007) suggest seven of them, but in the current research we will be focusing on the two particular: a) presence of a pause on an EDU border; b) presence of a nuclear accent inside an EDU [p. 57-60].

4) **Sample annotation**. Many books and manuals on corpus linguistics consider the issue of sample annotation. In the paper, the term "*annotation*" is understood as linguistic and extralinguistic data about the (corpus) sample, its parts and smaller units displayed in the form of tags [McEnery and Hardie, 2012, p. 29; Garside, Leech & McEnery, 2013, p. 2]. To *annotate/encode* means to add such information into the sample. Garside, Leech and McEnery (2013) notice that annotation should be performed according to a list of certain principles: a) verbal (raw) component should be easily separated from annotations; b) annotation should include metadata; c) the basis of annotation should be a well-established theory [p. 6-7]. Dealing with spoken discourse implies drawing the line between transcription and prosodic annotation. O'Keefe and McCarthy (2012) see this difference in the abstraction degree; annotation deals with data interpretation, while transcription is both about verbal-prosodic representation and interpretation of data; transcription always precedes annotation [see also Garside, Leech & McEnery, p. 2]. Linguistic annotation (assignment of tags to linguistic units) can be performed at any level including discourse, e. g. *part-of-speech* (PoS) annotation simultaneously relates to morphology and semantics (and, in a sense, syntax).

Results

It was previously noticed that transcription and morphosemantic annotation act as the most general and complex models within the study. The model of discursive transcription represents an adapted system developed by Kibrik and Podlesskaya (2007). As it displays twenty five phenomena of

different nature, we are going to abstract from the system as a whole and concentrate on its particular aspects that will help us to construct an image of the most typical (hence, relevant) EDU in the eSports commentary. Namely, these aspects are as follows: duration of EDUs, pauses, accents and nuclear accents, tone directions and illocutive meanings. Similarly, as morphosemantic annotation model contains PoSes of words (including multi-token ones) together with their lemmas, we will only focus on PoSes. Let us elaborate on transcriptional data.

The sample incorporates ***1819 EDUs*** and an average length of an EDU equals to ***4.1 tokens***, whereas its duration is ***1.26 seconds***. About 99.1 % of EDUs contain verbal components (i.e. words). There are ***989 pauses*** in the material; 69 % of them (683) are outer pauses; the rest 31 % (306) are inner ones. By dividing the number of outer pauses (683) by the number of EDU boundaries (1701), we get that pauses are present between 40.01 % of EDUs. In average, regardless of its type, a pause lasts ***0.324 ms***. About 84.7 % (838) of the pauses are absolute or respiratory. Inner pauses turned out to be respiratory in 74 % of cases; outer ones are respiratory in 74 % of instances. Filled and mixed pauses, being very rare, occur in 7.1 % and 8.1 % of cases correspondingly. Interestingly, they are distributed between outer and inner pauses in a very different manner: filled and mixed pauses have a tendency to occur inside EDUs (distribution is 4.3 % and 6.3 % for outer; 14 % and 11 % for inner). The outlined statistics are briefly summarized in the table 2 below.

Table 2. **Data on EDUs and pauses**

EDU length (avg)	4.1 tokens, 1.26 seconds		
EDU boundaries	present in 40.01 %; absent in 59.99 %		
Pause filling	**Pause location**		
	Outer	**Inner**	**Total**
Respiratory/Absolute	*89.3 %*	*74 %*	*84.7 %*
Filled	*4.3 %*	*14 %*	*7.1 %*
Mixed	*6.3 %*	*11 %*	*8.1 %*

Now we are going to consider the issue of accents. There are ***2439 accents*** in the material; ***1814 of them are nuclear***. Nuclear accent is present in 1672 EDUs (***92 %***), the rest of EDUs either contain several nuclear accents (3 %) or do not contain any (5 %). This may happen, for example, due to false starts or prosodic merging. With the help of special denotations and programs, we have modeled the location of the nuclear accent

in an EDU; the result of this procedure is presented in the table 3 below. In the scheme, 'C' stands for a so-called "common part", i. e. part without nuclear accents, whereas 'N' stands for a token with the nuclear accent. It can be observed that the nuclear accent gravitates towards the end of an EDU ('CN' and 'N' structures together occur in almost 80 % of cases). In 10.5 % of cases, the nuclear accent is located in the middle (between "common parts"). Other structures such as one with nuclear accent preposition ('NC') appear even rarer. Table 3 also contains several examples of EDUs that match the modeled structures.

Table 3. **Probabilistic model of nuclear accents location in EDUs**

Pattern	**'CN'**	**'CNC'**	**'N'**	**'C'**	**'NC'**	**'CNCN'**	**Other, e. g. 'NCNC', 'NNN', 'CNCNCN'**
Freq-cy	*1264*	*191*	*186*	*80*	*36*	*26*	*38 in total*
%	*(69.5 %)*	*(10.5 %)*	*(10.3 %)*	*(4.4 %)*	*(1.9 %)*	*(1.4 %)*	*each <1 %*
Example (Rus.)	*Потенциал \есть,*	*их /\много довольно,*	*\да,*	*То есть ==*	*\прямо / скажем.*	--	--
(Eng.)	*Potential \ exists,*	*them /\ many rather,*	*\yes,*	*Namely ==*	*\directly /saying.*	--	--

According to quantitative data on the directions of tones in accents presented in the table 4, eSports casters predominantly operate with ***simple tones*** such as "/" (rising) and "\" (descending). Pure rising and descending tones take place approximately in 84 % of both nuclear and other accents. Among other frequent tone directions, there are allophones of a rising tone: "/\" (rising-descending; almost 10 % of any accents) and "–" (straight; almost 4 % of any accents). These allophones are typically used when enumerating changing game events. As for descending tones, their "substitutes" such as "\/" (descending-rising) can be considered exotic, as well as all other tone directions in EDUs within the sample.

Situation with the illocutive meanings that EDUs sequences bear in the eSports commentary is simple and quite predictable: ***96 %*** of speech acts are ***statements***. Exclamations and questions almost entirely share the remaining 4 %; all other illocutions (including directives and complex illocutive meanings) seem to appear with a probability less than 1 %.

Table 4. **Statistics on tone directions**

Tone type	/ (rise)		\ (descent)	/\ (rise-descent)	– (straight)	\/ (descent-rise)	Other, e. g. \–, /–
Frequency (nuclear accent)	*891*		*630*	*176*	*73*	*16*	*22 in total*
Frequency (any accent)	*1215*		*845*	*241*	*95*	*18*	*25 in total*

Now let us briefly describe the basis of the ***part-of-speech annotation model***. PoS tags that constitute it were taken from three separate systems: a) *Universal Dependencies Treebank* (ADJ — adjective; ADP — preposition; ADV — adverb; AUX — auxiliary word; CCONJ — coordinating conjunction; SCONJ — subordinating conjunction/allied word; INTJ — interjection; NOUN — common noun; PROPN — proper noun; VERB — verb or infinitive; NUM — quantitative number/word; PART — particle or discursive marker; PRON — pronoun); b) *pymorphy2 Python library* (PRT — participle; GRND — adverbial participle; COMP — comparative, PRED — predicative word) and c) *discursive transcription system* (tags were suggested by us; NF — broken lexeme; UNK — illegible fragment). The necessity to combine these tags into one system is driven by the facts that: 1) rare forms for Russian spoken discourse (e. g. adverbial participles) bring certain meaning and are worth being specified; 2) words with unknown PoSes cannot be omitted due to the objectivity principle. As a result of the morphosemantic annotation procedure, all words in the sample (including multi-token ones such as *как раз*, eng. *just*) got their PoS tags. Then followed the calculation of relative frequencies for the patterns from three functional submodels of PoS model: unigrams, bigrams and EDU structures. The output was quantitative data that gives an idea of the morphological, semantic and syntactic nature of the structures frequently used by eSports casters.

Statistics on ***unigram patterns*** in the eSports commentary can be found in table 5 below. In brief, it illustrates the picture with the prevailing core part-of-speeches, (proper) nouns (20 % + 5.1 %) and verbs (17.2 %), typically forming nominative and predicative EDUs. Among auxiliary part-of-speeches, as we can see, particles & discursive markers (12.3 %) and prepositions (8.4 %) are the most numerous. Adverbs and adjectives, respectively, share 10.7 % and 6 %.

Table 5. **Unigram PoS patterns**

NOUN	VERB	PART	ADV	ADP	PRON	ADJ	CCONJ	PROPN	SCONJ
20 %	*17.2 %*	*12.3 %*	*10.7 %*	*8.4 %*	*7.2 %*	*6 %*	*5.2 %*	*5.1 %*	*2.3 %*
PRED	**NUM**	**NF**	**AUX**	**INTJ**	**UNK**	**COMP**		**PRT**	**GRND**
1.9 %	*1.4 %*	*0.6 %*	*0.6 %*	*0.4 %*	*0.2 %*	*0.2 %*		*0.1 %*	*0.1 %*

With respect to ***bigram part-of-speech patterns***, it was decided to take four most frequent structures for an illustration. They are presented below (examples are in square brackets):

Preposition + Noun, 363 occurrences. Example: *Попытка выйти* [*в \спину*], (*Attempt to_enter* [*from \rear*],);

Verb + Noun, 265 occurrences. Example: [*/Продолжают греки*] *убивать /курьеров,* ([*Continue Greeks*] *to_kill /couriers*);

Adjective + Noun, 256 occurrences. Example: [*невероятная игра*] *··(0.4) от /\обоих \игроков*. ([*unreal game*] *··(0.4) from /\both \players*);

Particle/discursive marker + Noun, 205 occurrences. Example: *следом* [*правда /\омнинайт,*] (*next* [*really /\omniknight,*]).

And finally, four the most frequent ***part-of-speech patterns of the EDUs*** in the eSports commentary are the following:

Participle/discursive marker, 69 occurrences. Example: *\да,* (*\yes,*);

Noun, 68 occurrences. Example: *\Дифуза,* (*\Diffusal,*);

Verb + Noun, 54 occurrences. Example: *Прыгнул \кентавр, (Jumped \centaur,*);

Noun + Verb, 36 occurrences. Example: *вард /стоит,* (*ward /dwells,*).

Discussion

Having represented the quantitative and qualitative data on the typical structural and semantic characteristics of EDUs in the eSports commentary, on the typical patterns within the genre, now it is time to comprehend and interpret the obtained results. There are two assumed sources of explanation: extralinguistic context (situation) and consciousness specificity.

1. The first group of phenomena includes ones that can be interpreted with appeal to Dota 2 commentators› necessity ***to speak in a fast pace*** and do it ***almost constantly without any preparation***. Dota 2 is a game where events can change dramatically within a few seconds, and commentators should be able to keep up when covering them in live. They simply do not have a right to keep silence for too long, as there is audience waiting for the

emotions expressing and explanations of what is happening. Apparently, it results in deviations from the norm of the Russian spoken discourse shown in Kibrik and Podlesskaya (2009): average EDUs are pronounced faster (1.26 s. for the commentary vs. 1.85 s. for the common discourse); pauses shift from the EDU boundaries inwards (40.01 % boundary pauses for the commentary vs. 62.7 % for the common discourse); during tense game moments everything except nouns (à objects) and verbs (à actions) is omitted for time saving; the most frequent EDUs patterns are very short (participles, nouns), etc.

2. The second group contains phenomena that correspond to the norm and seem to reflect the most general ***properties of human consciousness***. Whereas the speech apparatus can be trained, consciousness limits the volume of information that can be simultaneously kept active and requires time for switching between frame slots [Chafe, 1996]. Therefore, commentators resort to various pauses usage in almost similar way as children in Kibrik and Podlesskaya's (2009) study: both in casters' and children's speech absolute pauses occur in 84.5 % of cases, while filled and mixed pauses more often take place inside EDUs; discursive markers (particles) are often used for taking time to think. Nuclear accents are typically held (92 %) in order to denote the purpose of the current discursive step and focus audience's attention on the most significant aspects of the narrative. At the same time, accents gravitate towards postposition (which is typical for Russian spoken discourse), whereas tone directions in them are usually simple.

Conclusion

Summing up everything that has been said, through addressing functional language modeling and eSports sphere characteristics, the ideas of cognitive-discursive, structural and corpus linguistics, we managed to identify of the most relevant EDU patterns in the eSports commentary and to make certain assumptions about their role in the genre. Thus, with a high probability of more than 50 %, a random elementary discursive unit in the eSports commentators' speech: 1) is quickly pronounced; 2) is not separated from another EDU with an outer pause; 3) if contains a pause, it is short and absolute/respiratory; 4) embeds the nuclear accent with a simple tone direction in a postposition; 5) contains nouns and verbs. It is not, however, required that a typical EDU contains more than one such characteristic with a chance of 50 %; therefore, their dependencies should be studied in more detail.

The explanation for these and other results presented in the paper is assumed to consist in the extralinguistic factors, i. e. situation in which

commentators produce their speech, and human consciousness features. From the one side, the need to speak constantly and specificity of Dota 2 discipline require that eSports casters use their speaking skills at the maximum level. From the other side, human consciousness that controls language use limits commentators› capacities, forces them to resort to pauses and discursive markers. Aiming to keep up with such conditions and to achieve their professional goals, they train to use short and informative structures in their speech and strive not to load these structures with redundant elements. The latter, however, does not always work out.

Lastly, we would like to note that the findings are subject to an important constraint. Namely, the research sample is not fully representative and balanced, as its compilation is in process. When using the presented statistics in other research, e. g. in a more thorough investigation of the eSports commentary production, this fact should be taken into consideration.

References

Apresjan, Y. (1966). *Principles and methods in contemporary structural linguistics*. Moscow: Prosveshchenie Publ. (In Russian)

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

Baranov, A. (2001). *Introduction to applied linguistics*. Moscow: Editorial URSS Publ. (In Russian)

Big encyclopedic dictionary. Linguistics (1998). 2nd ed. Moscow: Bol'shaia rossiiskaia entsiklopediia Publ. (In Russian)

Boguslavskaya, V., Azizulova, A., Budnik, E., Sharakhina, L. (2018). Linguistic representation of ESports media community. *Sign. Problematic Field of Media Education*, 4 (30): 104–111. (In Russian)

Carlson, L., Marcu, D., Okurowski, M. E. (2003). Building a Discourse-Tagged Corpus in the Framework of Rhetorical Structure Theory. In: J. van Kuppevelt, R. Smith. (eds), *Current and New Directions in Discourse and Dialogue. Text, Speech and Language Technology*. Vol. 22 (pp. 85–112). Dordrecht: Springer. https://doi.org/10.1007/978-94-010-0019-2_5

Chafe, W. (1994). *Discourse, consciousness, and time*. Chicago: University of Chicago Press.

Du Bois J., Schuetze-Coburn S., Cumming S., Paolino D. (1993). Outline of discourse transcription. In: J. Edwards, M. Lampert (eds), *Talking data: Transcription and coding in discourse research* (pp. 45–87). Hillsdale, NJ: Lawrence Erlbaum.

Du Bois, J. (1992). Discourse transcription. *Santa Barbara papers in linguistics*, 4: 1–225.

Garside, R., Leech, G., McEnery, T. (eds) (2013). *Corpus Annotation: linguistic information from computer text corpora*. London; New York: Routledge.

Hamari, J., Sjöblom, M. (2017). What is eSports and why do people watch it? *Internet Research*, 27 (2): 211–232. https://doi.org/10.1108/IntR-04-2016-0085

Himik, V. (2015). Russian colloquial speech: general notion, training and terminology issues. In: *Selected works: proceedings of the 43th International Philological Conference, SPBU, March 2014, 11–16* (pp. 460–471). St. Petersburg: St. Petersburg University Press. (In Russian)

Kempe-Cook, L., Sher, S., Su, N. (2019). Behind the voices: The practice and challenges of esports masters. In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, May 2019* (pp. 1–12).

Kibrik, A. (2003). *Discourse analysis in cognitive*. Dr. Sci. thesis. Moscow: Institute of Linguistics of the Russian Academy of Sciences. Available at: https://www.dissercat.com/content/analiz-diskursa-v-kognitivnoi-perspektive (accessed: 09.11.2022). (In Russian)

Kibrik, A., Podlesskaya, V. (eds) (2009). *Stories about dreams*. Moscow: Iazyki slavianskikh kul'tur Publ. (In Russian)

Kravtsova, Y. (2014). Modeling in modern linguistics. *Vestnik of Zhitomir University*, 5 (77): 181–189. (In Russian)

Levelt, W. (1993). *Speaking. From intention to articulation*. Cambridge, MA: MIT Press. https://doi.org/10.7551/mitpress/6393.001.0001

Li, L., Uttarapong, J., Freeman, G., Wohn, D. (2020). Spontaneous, Yet Studious: Esports Commentators' Live Performance and Self-Presentation Practices. *Proceedings of the ACM on Human-Computer Interaction*, 4 (CSCW2): 1–25.

Mann, W., Thompson, S. (1988). Rhetorical structure theory: toward a functional theory of text organization. *Text & Talk*, 8: 243–281.

McEnery, T., Hardie, A. (2012). *Corpus Linguistics: Method, theory and practice*. Cambridge: Cambridge University Press.

Mikulinskiy, A. (2021). Comparative analysis of vocabulary of a gamer and eSports caster in Dota 2. *Theoretical and applied aspects of speech activity*, 7 (14): 89–98. (In Russian)

Mikulinskiy, A. (2021). *ESports commentary in Dota 2 sphere: means of expressiveness and corpus aspect*. Unpubl. grad. thesis. Nizhny Novgorod: HSE Publishing House. (In Russian)

O'Keefe, A., McCarthy, M. (eds) (2012). *The Routledge handbook of corpus linguistics*. London: Routledge.

Pankina, V., Khadieva, R. (2016). ESports as a phenomenon of the 21st century. Physical Culture. *Sport. Tourism. Motor Recreation*, 1 (3): 34–38. (In Russian)

Randhawa, N. (2015). The Games, the Audience, and the Performance. *ESports Yearbook 2013/14*, 14: 15–22.

Revzin, I. (1977). *Modern structural linguistics. Problems and methods*. Moscow: Nauka Publ. (In Russian)

Sukhodimtsev, P. (2016). ESports competition commentator: entertainment or profession? *Professional education and labor market*, 3: 16–18. (In Russian)

Sukhodimtseva, A., Sukhodimtsev, P. (2016). Profession of ESports commentator as a social phenomenon. *International Scientific Review*, 11 (21): 93–95. (In Russian)

Tomlin, R., Forrest, L., Ming, M., Hee, M. (1997). Discourse semantics. In: van Dijk (ed.), *Discourse as structure and process* (pp. 63–111).

Van Dijk, T., Kintsch, W. (1983). *Strategies of Discourse Comprehension*. New York: Academic Press.

Wagner, M. (2006). On the Scientific Relevance of eSports. In: *International conference on internet computing*. Danube University Krems, June 2006 (pp. 437–442).

Zaripov, A. (2016). Lexical aspect of the eSports discourse. *Philology Sciences. Theoretical and Practical Issues*, 2–1 (56): 99–101. (In Russian)