Investigating the Differences Between Prepared and Spontaneous Speech Characteristics: Descriptive Approach

Galiya M. Gatiyatullina[1,*], Ludmila A. Gorodetskaya[2], Marina I. Solnyshkina[1] and Elzara V. Gafiyatova[1]

[1]*Department of Theory and Practice of Teaching Foreign Languages, Institute of Philology and Intercultural Communication, Kazan Federal University, Kazan, Tatarstan, Russia*

[2]*Department of Theory of Teaching Foreign Languages, Moscow State University, Moscow State University, Moscow, Russia*

Abstract: In the modern EFL paradigm, pre-task planning time is viewed as a norm. Pre-task planning time is one of the central concerns of teachers, test-developers, as well as researchers. Pre-task planning is planning a speech before performing a task, and it also involves rehearsal and strategic planning. The paper addresses the problem of pre-task planning advisability for A2 Russian EFL speakers. The research presented in this paper examines the structure, breakdown, repair, syntactic complexity, lexical diversity as well as the accuracy of the discourse produced by 145 Russian participants of the English language competition held in Kazan, Russia, in January 2020. The discourse analysis revealed that the pre-task time is used by A2 EFL speakers not only to focus on a dialog but also to elicit a topic text from memory, thus focusing on form rather than meaning. Hence, in A2 tests prioritizing meaning over form and measuring the ability for spontaneous speech, the one-minute pre-task planning time is viewed as questionable.

Keywords: Oral performance, pre-task planning, A2, Russian EFL speakers discourse.

INTRODUCTION

Pre-task planning time is one of the central concerns of teachers, test-developers, and researchers. Unfortunately, the results of the research conducted in psychology, cognitive science, psycholinguistics, and discourse analysis are too inconsistent with providing practitioners with a theoretical foundation for calculating the time ratio of preparation and oral performance (Abdi, & Basarati, 2018).

In the modern EFL paradigm, pre-task planning time is viewed as a norm (Mann & Taylore-Knowles, 2006). The time provided to Russian students before oral performance is equal or little less than speaking time and ranges between one and ten minutes (Verbickaya *et al*., 2015). The following three interconnected postulates support the idea: 1) humans possess a limited capacity to process information (McLaughlin *et al*., 1983; Posner & Klein, 1973; Solovyev *et al*., 2019); 2) students'/examinees' attention can be profitably channeled through making instructional choices (Schmidt, 1990); 3) L2 speakers' attention to one area of the language (i.e., form) is typically drawn at the expense of another (i.e., content) (Foster & Skehan, 1999). Teachers and test-developers view pre-task planning as a possibility for test-takers to focus on what and how to speak during planning time so that they can focus mainly on meaning during the oral performance. Thus, these theories a priori implicate that planning before speaking contribute to control the level of cognitive demand imposed by potentially unfamiliar topics and establish a fair environment for test-takers. In this regard, the question of the rationale for pre-task planning time and its amount is, in fact, a question of attention between the form and meaning (Nitta & Nakatsuhara, 2014).

Before the introduction of EGE, i.e., the Unified State Exam in English in the Russian Federation in 2007, the choice had predominately been made for the form: examinees' performances were mainly rated by grammar and vocabulary accuracy only (Solnyshkina *et al*., 2014; Keaton, & Giles, 2016). The present reality, when educators in Russia are looking for tools to assess real-life performance, not reproduction mechanisms, poses the problem of pre-task planning practicality. This problem becomes more urgent with A2 speakers who tend to use planning time for rehearsals of the topics memorized before (Abdi, & Basarati, 2018).

The study has been conducted to verify the following hypotheses: (1) 'With pre-task time provided, A2 speakers tend to reproduce rather than produce speech' and (2)'The length of planning time affects the discourse produced', the study focuses on whether A2 test-takers should be provided with planning time before the oral performance. But acknowledging the

*Address correspondence to this author at the Department of Theory and Practice of Teaching Foreign Languages, Institute of Philology and Intercultural Communication, Kazan Federal University, Kazan, Tatarstan, Russia;
Tel: +79270409055; E-mail: mingaz.ll@yahoo.com

E-ISSN: 1929-4409/20

partiality of the topic chosen and the limitation of the data, in this article, the authors try to avoid any predictions but instead raise the problems.

LITERATURE REVIEW

Linguists (Ellis, 2005) distinguish task-based planning consisting of two principal types: pre-task planning and within-task planning. Pre-task planning is planning a speech before performing a task, and it involves rehearsal and strategic planning. Rehearsal or focus-on-form is the repetition of speech before the performance, while in strategic planning, a speaker focuses on the content of speech to deliver. Within-task planning can be either pressured or unpressured, depending on whether a concise time limit is imposed. With a pressured time limit or on-line planning, a speaker produces spontaneous or unprepared speech. It is traditionally accepted that "constraints in attentional capacity during task performance result in one aspect of performance being prioritized and improves, whereas another aspect receives less attention and remains the same" (Foster & Skehan, 1999; Yuan & Ellis, 2003). Such phenomena known as trade-off effects have been investigated intensively, mostly in dialogue tasks (Nitta & Nakatsuhara, 2014), but research in the area target mostly B2 – C1 speakers and view pre-tasks planning time as necessary to regulate the cognitive demand imposed by potentially unfamiliar speaking topics thus improving test-takers' oral performance (Nitta & Nakatsuhara, 2014).

As for comparative studies of pre-task planning effects on A2 speakers' fluency, accuracy, and complexity in monologic performances, it is, to the best of our knowledge, still an under-investigated element in EFL and language assessment.

The results of the previous studies also contain partially contradicting and, in rare cases, inconsistent findings: Mehrang & Rahimpour (Mehrang & Rahimpour, 2012) report that planning time did not affect the accuracy and fluency of the learners' performances. While Yuan & Ellis (Yuan & Ellis, 2003) argue that pre-task planning impacts language production positively, especially where the focus is on fluency and complexity. Ahangari and Abdi (Ahangari & Abdi, 2011) demonstrate that learners with the opportunity to plan before task performance may produce more complex language, although no positive effect is evident in the accuracy of learners' oral performance. Some benefits of pre-task planning are also reported by Tavakoli (Tavakoli & Skehan, 2005) and Wigglesworth (Wigglesworth, 1997). While Wigglesworth in later, his studies (Wigglesworth & Elder, 2010) revealed limited or no effects. In our previous studies (Solnyshkina *et al*., 2016) we found that speakers rehearsed rather than planned their speech with their monologues with a smooth start and even tempo in the beginning and proceeding to quite frequent hesitations and false starts from the third or the fourth utterance. In this study, we investigated dialog speech, which speakers were asked to make up after a minute break.

Galyashina (2003) suggests the following scale of spontaneity of speech:

1) spontaneous speech (spontaneous dialogue, spontaneous monologue);

2) semi-spontaneous speech (interview, story on a given topic, reproduction of someone else's speech, deliberate speech according to a pre-compiled plan, stereotyped speech on a template text); prepared speech (retelling and reading aloud) (Galyashina, 2003).

METHODS

Semi-spontaneous and spontaneous types of speech are of particular interest, since the speakers most often resort to their favorite speech patterns, whereas the process of reading is influenced by graphs, reading techniques, the speaker's familiarity with the text, visual acuity, illumination, and other extralinguistic factors. Physical variables in the semi-spontaneous speech are of less importance, however psycho-emotional variables are significant. Spontaneous speech due to its unpreparedness and situational conditioning is characterized by a number of distinctive features, which to a certain extent can be regarded as deviations from the norms of prepared speech. One of the characteristics of spontaneous and quasi-spontaneous conversational speech is affectivity or emotional intensity. Previous research indicates that emotionality refers to «the expression of feelings, moods, subjective relationships» (Ahmanova, 2004: 211).

This study was designed to address the following research questions:

1. What are typical A2 discourse features (structure, breakdown, repair, syntactic complexity, lexical diversity, accuracy) in test-takers' oral performance after a minute pre-task planning time?

2. What are typical A2 discourse features (structure, breakdown, repair, syntactic complexity, lexical diversity, accuracy) in test-takers' oral performance in a spontaneous dialog?

3. Is a minute pre-task planning time advisable for A2 EFL test-takers?

'Read speech' and 'spontaneous speech' are terms broadly used to refer to two typically contrasted speech styles, though the terms themselves do not refer to any fixed set of conventions or inherent set of prosodic features. Nevertheless, read and spontaneous speech are often described in terms of their differences, at syntactic, lexical and prosodic levels, and are suggested to be perceptually distinguishable based on prosody alone. A typical conceptual contrast made between these two styles is of spontaneous speech as informal, dynamic and unrehearsed (as from a casual conversation) versus read speech as scripted and formal (as from a news reader) (Laan,1997:45).

The Learner's corpus has 21628 tokens and consists of monologue speech (11623 tokens) and dialog (10005 tokens). The speech was recorded from 145 participants who were secondary school students aged 11 – 13 participating in the English language competition in Kazan, Russia, in January 2020. The English language proficiency level of the participants was assessed as A2 CEFR based on the results of the proficiency test (http://www.stgiles-international.com). The test contained grammar and vocabulary questions along with Reading, Listening, and Writing tasks.

The speaking tasks developed by the local Department of Education were formally in full concordance with the Manual "Relating Language Examinations to the Common European Framework of Reference for Languages: Learning, Teaching, Assessment (2009) (North *et al*., 2009) issued by the Council of Europe: the competitors were expected to demonstrate the ability to "give a simple description or presentation of people, living or working conditions, daily routines, likes/dislikes etc. as a short series of simple phrases and sentences linked into a list". The competition organizers expected each participant to produce a two-minute oral monologue elicited utilizing a card with one of the two tasks written on it:

Task A: You have one minute to prepare a talk on your favorite kind of sport. Do you enjoy watching or playing? Who is your favorite sportsman?

Task B: You have one minute to prepare your talk on your favorite dish. What food do you prefer? Can you cook something by yourself?

Note-taking was prohibited as the pre-task planning was expected to be used for strategic unguided planning.

Shortly after a monologue, the participants were offered a minute break and a task to participate in a dialog with another participant. The participants were informed of a dialog topic right before their two-minute break.

Task 2. Your friend is having a birthday party. Discuss with a partner what is the best birthday present: money, a book, a computer game or a fashionable T-shirt.

Participants' performance was assessed by two raters who applied the following analytic approach to score language features: Task response (max 10 points), Coherence /cohesion (max 10 points), Grammar range and Accuracy (max 3 points), Lexical Resource (max 3 points), Fluency Pronunciation (max 2 points), Time (max 2 points). The participants were informed about the rating scale. The speaking time fell within the range 0 – 10.4 seconds.

In the study discourse analysis, conversation analysis (CA) was used. To calculate the Measure of Textual Lexical Diversity (MTLD) (McCarthy & Jarvis, 2010), the Textinspector (https://textinspector.com) was used.

As systems become more habitable and allow users to speak naturally, speech recognizers and parsers are going to have to deal with events not present in written text or read speech. Spontaneous speech contains a number of phenomena that cause problems for current systems.

- filled pauses - noises made by the speaker that don't correspond to words (ah, uh, um, etc).
- restarts - repeating a word or phrase. The original word or phrase may be complete or truncated.
- interjections - extraneous phrases as in "on line thirty, I guess it is".
- unknown or mispronounced words
- ellipsis

- ungrammatical constructions - Users make errors of agreement (sub-verb, number, etc) and may use constituents in unusual orders ("to the utilities cell add fifty dollars").

These phenomena violate constraints currently used by speech recognizers to increase performance. This can cause complete recognition failure for an utterance (Ward, 2009: 137). In the context of speech technology research, the term 'spontaneous-speech' has traditionally been used in contrast to 'read-speech' as an indicator of the degree of control in speech utterance production, according to whether the speech content is generated in real-time, while speaking, or is simply converted from text through a process of reading. Spontaneous speech is thought to be 'more noisy' (hesitations andfillers) and 'less-well-formed' (or un-grammatical). However, it is an over-simplification to assume that spontaneity is a binary attribute of speech. There are degrees of spontaneity, even in read speech, and speaking styles can vary in a range between the highly-rehearsed formal presentation style (e.g., for broadcasting and public-speaking), and the intimate chatting betweeen friends and family members. With the former, the controlled structure of the speech arises from a predominance of lexical information (and often by a reliance on a written text as the original basis for the speech), but with the latter, the degree of shared common-knowledge is much higher, and much of the spoken interaction takes place in a non-verbal form. Often its purpose is not to impart information, but simply to be social. The JST/CREST ESP Corpus exemplifies the latter. It consists of wholly unprepared speech, with controls for the degree of familiarity between speaker and hearer. In this paper, we present results of an analysis of part of this corpus, showing that the same lexical string, spoken by the same speaker, often carries different paralinguistic information, and we confirm that independent listeners can form a similar context-independent interpretation of this 'meaning-behind-the-words' from similarities in the prosodic and voice-quality parameters (Campbell, 2017).

RESULTS AND DISCUSSION

Communication is an important aspect of quality of life, and speech is the primary means of human communication. In adults with intellectual disabilities (ID), speech communication is often troubled by disordered speech production and/or impaired hearing , resulting in miscommunication and consequently impairing social interactions, possibly behavioural problems and isolation (Roberts & Malkin, 2007:30).

When speech understanding systems are used in real applications, they will have to deal with phenomena peculiar to spontaneous speech. People use language differently when they speak than when they write. Spoken language contains many interjections, filled pauses, etc. Speakers often don't use well-formed sentences. They speak in phrases, have restarts, etc. Systems designed for written or read text will encounter serious difficulties processing such input (Ward, 2009: 137).

As in our previous research (Solnyshkina *et al.*, 2016), the participants' textual products were audio-recorded and transcribed. Pauses were referred to as hesitation, if silence lasted between 0.3 to 0.4 sec., or unfilled if silence was more than 0.5sec. All pause fillers, such as *um, ah, ham, er,* were measured with a stopwatch and registered in the transcripts. The following notations were used in the transcripts: three dots represent a silent pause, Russian inclusions were transcribed in Latin graphics in square brackets, e.g. [nu], [tam], [tak], [da], etc.

The transcripts were then analyzed with the use of discourse analytic measures adapted from (Nitta & Nakatsuhara, 2014): fluency (speed – the number of words per second, breakdown – the number of lexicalized / unlexicalized pauses per speaking time, repairs – the number of repetitions and corrections), complexity (syntactic complexity – the number of clauses, lexical diversity), accuracy the number of errors per 100 words.

As it was anticipated, all participants as winners of school and district English language competitions demonstrated the ability "to use a series of phrases and sentences to describe in simple terms family and other people, living conditions, educational background, and present or most recent job" (CEFR Level A2) (ttp://www.stgiles-international.com), and generally produced clear, coherent discourse using several cohesive devices. However, the performances were not utterly homogeneous across the population: some participants achieved a B1 standard, while others managed to A2 band only.

Spontaneous speech has another dimension of difficulty for automatic processing when more than one speaker is involved. As described in classic work on conversation analysis (Abdi, Basarati, 2018), turntaking

involves intricate timing. In particular, speakers do not alternate sequentially in their contributions as often suggested by the written rendition of dialog. Rather, listeners project the end of a current speaker's turn using syntax, semantics, pragmatics, and prosody, and often begin speaking before the current talker is finished. It is also interesting that familiarity with the other talker does not appear to affect the rate of overlap, since there is no difference in overall rates between the CallHome and Switchboard conversations. Modeling realistic turn-taking has a fascinating application to dialog system design; some researchers are developing conversational systems that can mimic human turn-taking by providing backchannel responses, e.g. Overlap in turn-taking also introduces several problems for many current offline automatic speech processing tasks. An obvious difficulty is the acoustic modeling of simultaneous speakers on a single recording channel. Relatively little work has focused on this problem to date, although source separation and auditory scene analysis techniques may ultimately lead to solutions. In this paper we focus on the impact of overlap for higher-level phenomena (Hirschberg, 2002).

Speech characteristics have not been studied widely in adults with ID in general, but Roberts *et al*. documented the following characteristics of the speech of adults with Down syndrome: consonant cluster reduction; final consonant deletion; unstressed syllable omission, mostly at the start of a multisyllabic word; and consonant substitution (e.g., a fricative sound /s/ becomes a plosive /t/). Errors in the speech of persons with Down syndrome have been characterised to be inconsistent, and the production errors and articulation difficulties observed mainly occur in the phonemes that are typically acquired in the final stages of normal speech-language development. Van Borsel attributed the errors to a developmental delay. Unfortunately, no data are available for adults with ID of mixed aetiology or other groups of persons with a specific ID (Kent & Vorperian, 2013: 179).

As mentioned above, the participants were asked to discuss with another participant about the best present for a friend (money, a book, a computer game or a fashionable T-shirt).

In dialogs, a hesitance part started after a participant takes his/her turn. While preparing for a dialog the participants were heard rehearsing their speech murmuring "*I think that... because...*". It is not a surprise that while speaking each participant tended to prove their opinion about each present. This part was produced or rather reproduced very smoothly without hesitations or false starts. In the beginning of the performance many of the participants demonstrated smooth fluency, syntactic complexity and lexical diversity associated with B1 band. Those were typically accurate stretches of speech produced at even tempo.

Most of such turns started with authorization discourse markers like: "*In my opinion*", "*I think*", as well as ended with discourse markers indicating giving the floor to a counterpart: "*What about t-shirt?*", "*what do you think*", "*And you*?", "*how about computer games*?" which make us think of them as cliché phrases.

Participants showed most of hesitation and repair while answering the questions of their counterparts. For example, in the following turn exchange, it is clear that a participant [B21] is not expecting to take the floor at that moment and started recollecting his own planned speech. After another short hesitation, he supports his opinion with an argument that changes into another

Table 1: Discourse Analysis

Discourse features		Reproductive/Planned phase	Hesitant phase	Productive phase
Duration, mean (sec)		6	4	43
Utterance Fluency	speed (words / second), mean	1,9	0,2	1,2
	breakdown (filled / silent pauses / speaking time), mean	0,17	0,6	0,3
	repairs (repetitions and corrections), mean	2, 1	10	4, 7
Complexity	syntactic complexity (the number of clauses)	8	0.2	5
	lexical diversity	32,6	6	28,8
Accuracy	accuracy (errors / 100 words)	3	37	20

hesitation because the participant [B 36] does not expect a new topic introduction.

In many formal written languages, punctuation is rendered explicitly. But spoken language is a stream of words, with no overt lexical marking of the punctuation itself. Instead, phrasing is conveyed through other means, including prosody. Sentence boundaries and other types of punctuation are useful for many types of automatic downstream processing (including parsing, information extraction, dialog act modeling, summarization, and translation), as well as for human readability. These methods are typically trained on text data, which contains punctuation. Modeling sentence-level punctuation can also improve speech recognition performance itself. Historically, speech recognition researchers have built language models based on sentences as found in text and then tried to acoustically segment the speech into sentence-like units. This is typically done by chopping at longer pauses and speaker changes. Pauses are relatively easy to detect and minimize the risk of fragmenting words in the process. Speaker changes are also often available, especially if speakers are recorded on different channels. For some applications, if a speaker produces one sentence at a time (for example, to a dialog system) there is typically little problem. But for processing of natural conversation, finding the sentence boundaries by machine is a challenge. Pauses are neither necessary nor sufficient indicators of sentence boundaries. People often string together sentences without pauses. And conversely, people pause (as during hesitations or disfluencies) at locations other than sentence boundaries. Computational models for finding sentence boundaries in speech typically involve a combination of N-gram language models (over words and boundary labels) and prosodic classifiers. Knowledge sources are often combined using an HMM framework. More recently, other model types have been used successfully, such as maximum entropy models and conditional random fields. Prosodic models have used probabilistic classifiers such as decision trees or neural networks, and can be improved by sampling and ensemble techniques. Additional approaches and features are described in. While initial research used hand-transcribed words as input, more recent work has studied the problems arising from imperfect recognizer hypotheses. In particular, work on strategies for using multiple recognition hypotheses appears promising (Shriberg2005).

Excerpt (1)

- [B 36] Ah, I'll give him some money because it is er it is very simple to present. What do you think?
- [B 21] But I think... the book will be the best present for him. Because er... he can er read it in the summer days when he can do anything.
- [B 36] But uhm it is very uhm challenging to choose a book for him. Do you know what books does he like?
- –[B 21] He likes er books about adventures like Treasure Island er and many other adventure books.

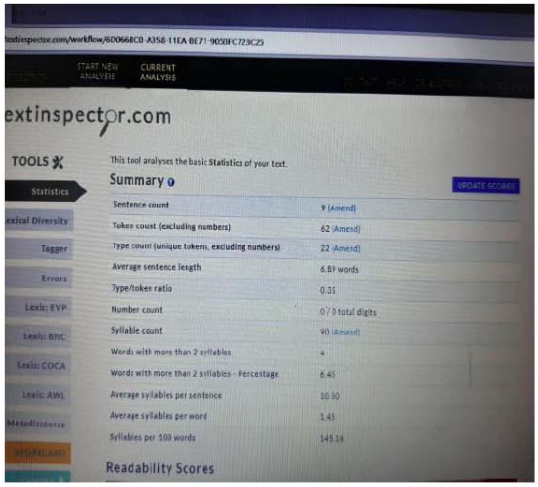
Two participants started their dialog with a long turn, which lasted up to 42 seconds, bringing it close to a monologue. The turn changed into another planned monologue given by a counterpart. The rest lost fluency, hesitated, and repeated phrases and words. The hesitant part was, on average, six words. In the Hesitant Phase, the number of stops, pauses, false starts, and restarts increased dramatically, speed reduced, accuracy and complexity dropped.

Syntactic complexity reduced a lot comparing to the delivered monologues. As mentioned above, most of the complex sentences were elaborated using *I think... Do you know.* Clauses or with *because* adverbial clauses.

The performances were also inconsistent in lexical diversity across the parts. The mean of “the range of different words” (McCarthy & Jarvis, 2010) analyzed with Textinspector (textinspector.com/workflow/) proved to be 32,6 words in the Planned Part, 15,08 words – in the Hesitant Part, and 28,8 words – in the Productive.

Figure **1** demonstrates the screen of a dialog analysis using the Textinspector website

The Productive Part started when apprehensions were overcome, new ideas began being formulated, the number of pauses (both lexicalized and non-lexicalized), repair and hesitations increased, the speed gradually increased, accuracy and complexity reduced. The participants began generating extemporaneous speech and focusing on the task and cohesion. The Productive part was, on average, 28,8



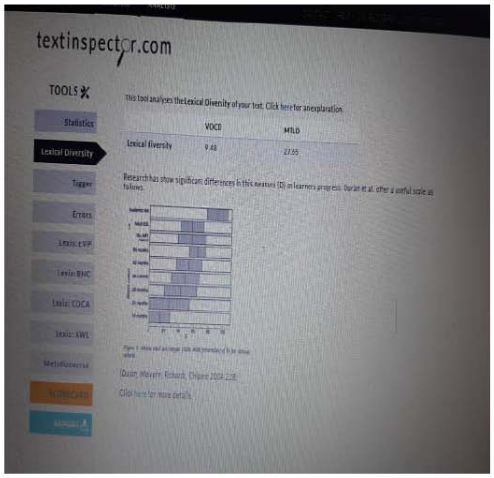


Figure 1: Lexical diversity analysis (Textinspector).

words. Another finding was that eye-contact was used more frequently than in the previous two parts.

CONCLUSION

The discourse analysis of A2 EFL test-takers oral performances revealed that the pre-task time is used by A2 EFL speakers not to plan a response but elicit a topic text from memory, thus focusing on form rather than meaning. It is mostly seen in dialogs when it takes several seconds for a participant to answer a counterpart's question. Hence, in A2 tests prioritizing meaning over form and measuring the ability for spontaneous speech, the one-minute pre-task planning time is viewed as not helpful. The speakers spent the time rehearsing their prases, i.e., focusing on form rather than content. Therefore, the turns containing rehearsed phrases were uttered smoothly with even tempo. The turns with spontaneous parts were filled with hesitations, repetitions, false starts. As a rule, the spontaneous part was at the beginning of a turn, when a speaker started answering a counterpart's question. The study of pre-task planning time offers new insights into potential differences between the ways A2 and higher proficiency level students use their pre-task time.

ACKNOWLEDGEMENT

The authors would like to thank Kazan Federal University for assistance in publishing the paper. The work is performed according to the Russian Government Program of Competitive Growth of Kazan Federal University. The research presented in Parts Discussion and Conclusion of the article was financially supported by the Russian Science Foundation, grant № 18-18-00436.

Appendix 2

- [B 30] Ah. I think. I think the best birthday present for you is a book. Because maybe you love reading...love...because maybe you uhm love reading books - [B 15] You are right. But reading books is pleasant, and I'll give my friend an English book. And what is the best present for you? - [B 30] I think the best present is a computer game because I... because I like playing computer games so it is very it is very uhm cool and how about a fashionable t-shirt? - [B 15] In my opinion, a fashionable t-shirt is a [unclear] present for a birthday. And what do you think about money? Is it a good present? - [B 30] I think it is a good present for my friend because ah you can buy ah you can buy something for money. Ah uhm er okay. That's all - [B 15] That's all.	– [B 21] Whoa. You know. What do you think about a present for our friend's birthday? –[B 36] Ah, I'll give him some money because it is er it is very simple present. What do you think? –[B 21] But I think... the book will be the best present for him. Because er... he can er read it in the summer days when he can do anything. –[B 36] But uhm it is very uhm difficult to choose a book for him. Do you know what books does he like? –[B 21] He likes er books about adventures like Treasure Island er and many other adventure books. –[B 36] Uhm but we can choose him a fashionable t-shirt and ah he will wear it to the school. –[B 21] No, I think the book will be the best present. –[B 36] Er, and what about a computer game? I think he likes playing a computer game. – [B 21]Well, let's buy a computer game. –[B 36] Okay.

REFERENCES

Abdi, R., Basarati, A. (2018). Legitimation in Discourse and Communication Revisited: A Critical View towards Legitimizing Identities in Communication. *International Journal of Society, Culture & Language.* 6(1), 86-100

Ahangari, S., & Abdi, M. (2011). *The Effect of Pre-Task Planning on the Accuracy and Complexity of Iranian EFL Learners Oral Performance*. Procedia – Socialand Behavioral Sciences. Elsevier. https://doi.org/10.1016/j.sbspro.2011.11.445

Ahmanova, O. (2004), Slovar lingvisticheskih terminov [Dictionary of linguistic terms], Moscow: Editorial URSS [in Russian].

CEFR Level A2. URL: http://www.stgiles-international.com/app/webroot/docs/Level-A2-Learner-Outcomes.pdf.

Elder, C., & Iwashita, N. (2005). *Planning for Test Performance: Does it Make a Difference?* In R. Ellis (Ed.), Planning and Task Performance in a Second Language (pp. 219−238). Philadelphia: John Benjamins. https://doi.org/10.1075/lllt.11.14eld

Elizabeth Shriberg (2005), Spontaneous Speech: How People Really Talk and Why Engineers Should Care, Speech Technology and Research Laboratory, SRI International, Menlo Park, CA 94025, USA International Computer Science Institute, Berkeley, CA 94704, USA, ees@speech.sri.com.

Ellis, R. (2005). Planning and task performance in a second language. *Philadelphia: John Benjamins Publishing Co*. 12-58. https://doi.org/10.1075/lllt.11

Foster, P., & Skehan, P. (1999). The Influence of Source of Planning and Focus of Planning on Task-Based Performance. *Language Teaching Research*, 3, 21-247. https://doi.org/10.1177/136216889900300303

Galyashina, E. (2003), Theory and practice of linguistic, analysis of media texts in forensic examinations and information disputes]. Materials of scientific-practical seminar,Moscow: Gallery [in Russian].

J. Hirschberg (2002), Communication and prosody. Speech Communication, 36. https://doi.org/10.1016/S0167-6393(01)00024-3

Keaton, S.A., Giles, H. (2016). Subjective Health: The Roles of Communication, Language, Aging, Stereotypes, and Culture. *International Journal of Society, Culture & Language*. 4(2), 1-10

Kent RD, Vorperian HK(2013), Speech impairment in Down syndrome: a review. J Speech Lang Hear Res 2013. https://doi.org/10.1044/1092-4388(2012/12-0148)

Laan, G. (1997), The contribution of intonation, segmental durations, and spectral features to the perception of a spontaneous and a read speaking style, Speech Communication, 22. https://doi.org/10.1016/S0167-6393(97)00012-5

Mann, M., & Taylore-Knowles, S. (2006). *Macmillan Exam Skills for Russia: Speaking and Listening*. Oxford: Macmillan.

McCarthy, P. M., & Jarvis, S. (2010). MTLD, vocd-D, and HD-D: A validation study of sophisticated approaches to lexical diversity assessment. *Behavior research methods*, *42*(2), 381-392. https://doi.org/10.3758/BRM.42.2.381

McLaughlin, B., Rossman, T., & McLeod, B. (1983). *Second language learning: an information-processing perspective*. *Language Learning, 33*,135—57. https://doi.org/10.1111/j.1467-1770.1983.tb00532.x

Mehrang, F., & Rahimpour, M. (2012). *The Impact Of Task Structure And Planning Conditions On Oral Performance Of EFL Learners.*

Nick Campbell(2017), Voice characteristics of spontaneous speech, ATR Human Information Science Laboratories, Keihanna Science City, Kyoto, Japan.

Nitta, R., & Nakatsuhara, F. (2014). *A Multifaceted Approach to Investigating Pre-Task Planning Effects on Paired Oral Test Performance, Language Testing, 31*(2), 147-175. https://doi.org/10.1177/0265532213514401

North, B., Figueras, N., Takala, S., Van Avermaet, P., & Verhelst, N. (2009). Relating language examinations to the Common European Framework of Reference for Languages: Learning, teaching, assessment (CEFR). *A Manual. Strasbourg: Council of Europe*.

Posner, M., & Klein, R. (1973). *On the Functions of Consciousness.* In S. Kornblum (ed.): Attention and Performance IV. London: Academic Press.

Roberts JE, Price J, Malkin C (2007), Language and communication development in Down syndrome. Ment Retard Dev Disabil Res Rev 2007;13. https://doi.org/10.1002/mrdd.20136

Schmidt, R. (1990). *The Role of Consciousness in Second Language Learning*. Applied Linguistics, *11*, 129–158. https://doi.org/10.1093/applin/11.2.129

Solnyshkina, M. I., Ziganshina, C. R., Sharifullina, E. A., & Gatiyatullina, G. M. (2016). The Effect of Pre-Task Strategic Planning on Russian A2 EFL Learners' Monologic Oral Performance. *International Journal of Humanities and Cultural Studies (IJHCS) ISSN 2356-5926*, *3*(2). https://doi.org/10.5539/elt.v7n12p1

Solnyshkina, M.I., Harkova, E.V., and A.S.Kiselnikov. (2014). Unified (Russian) State Exam in English: Reading Comprehension Tasks. *English Language Teaching, 7*(12), 19 November.

Solovyev, V., Solnyshkina, M., Ivanov, V., & Batyrshin, I. (2019). Prediction of reading difficulty in Russian academic texts. *Journal of Intelligent & Fuzzy Systems*, *36*(5), 4553-4563. https://doi.org/10.3233/JIFS-179007

Tavakoli, P., & Skehan, P. (2005). *Strategic planning, task structure, and performance testing*. In R. Ellis (Ed.), Planning and Task Performance in a Second Language (pp. 239−273). Philadelphia: John Benjamins. https://doi.org/10.1075/lllt.11.15tav

Textinspector, URL: http://www.textinspector.com/workflow/ 2FE5E7C6-5EDC-11E6-91DB-973EAFCE53D3.

Verbickaya, M., Mann, M., & Taylore-Knowles, S. (2015). *Testy dlya podgotovki k GIA in English*. Macmillan.

Wayne Ward (1988), Understanding spontaneous speech, Conference: Proceedings of the workshop on Speech and Natural Language. https://doi.org/10.3115/100964.100975

Wigglesworth, G. (1997). An Investigation of Planning Time and Proficiency Level on Oral Test Discourse. *Language Testing, 14*, 85−106. https://doi.org/10.1177/026553229701400105

Wigglesworth, G., & Elder, C. (2010). An Investigation of Effectiveness and Validity of Planning Time in Speaking Test Tasks. *Language Assessment Quarterly, 7,* 1−24. https://doi.org/10.1080/15434300903031779

Yuan, F., & Ellis, R. (2003). The Effects of Pre-Task Planning And on-Line Planning on Fluency, Complexity and Accuracy in L2 Monologic Oral Production. *Applied Linguistics, 24,* 1−27. https://doi.org/10.1093/applin/24.1.1

Received on 24-10-2020 Accepted on 29-11-2020 Published on 31-12-2020

DOI: https://doi.org/10.6000/1929-4409.2020.09.319