

Lexical Complexity of Learner Discourse: Interpersonal and Presentational Mode Descriptions in Russian

Julia Mikhailova

Introduction

In this case study I will analyze the lexical complexity of oral descriptions produced by learners of Russian in two different but related kinds of oral proficiency interviews, the Oral Proficiency Interview (OPI) and the Simulated Oral Proficiency Interview (SOPI). The OPI is a face-to-face or telephone interview of a speaker of a foreign language conducted by a tester certified by the American Council on the Teaching of Foreign Languages (ACTFL). In the OPI, description is elicited by the examiner in the context of live interaction. The description is thus in interpersonal mode, as defined by the National Standards for Foreign Language Learning, because the testee and tester interact as they negotiate the communication of information in the testee's responses to the tester's questions. The SOPI is a test administered by the Center for Applied Linguistics (CAL) in which the student being tested is given a test booklet and a tape with oral prompts. The test booklet contains instructions and visual prompts for each task. The student then records his or her speech in response to the oral prompts; students' responses must fit in the predetermined time allotted each prompt. There is no live interaction component in the SOPI and therefore speech elicited by the SOPI would be classified as presentational in mode according to the definitions of the National Standards for Foreign Language Learning. Raters certified by CAL receive the tape and assess the speech production for each of the prompts in accordance with criteria corresponding to the ACTFL Oral Proficiency Guidelines. National Standards for Foreign Language Learning distinguish between interpersonal and presentational modes of communication; therefore, the two testing formats may actually be assessing two different kinds of communication.

I analyzed samples of description in existing officially recorded oral proficiency interviews of students of Russian obtained through ACTFL (the OPI data), the Center for Applied Linguistics (the SOPI data), and oral exams (both

OPIs and SOPIs) taken by students of Russian at the Middlebury Russian School in the summer of 2003. I chose to investigate description for two reasons: First, description is one of the core functions required for advanced level speech according to the ACTFL Oral Proficiency Guidelines. Second, while previous research has examined other speech functions in proficiency testing (such as narration, giving instructions, stating opinion, and so on), description has not yet been considered either in Russian or across languages. Nor has there been a substantial comparison of interpersonal and presentational modes of speaking. Moreover, little work has been done in the area of analysis of testing for Russian. My research pioneers analysis in how modality affects the assessment of learning outcomes in the two predominant speaking foreign language oral proficiency exams [for any foreign language].

The focus of the study is Intermediate High level because previous analysis of American students' foreign Russian language proficiency (Brecht, Davidson and Ginsberg, 1993; Carroll, 1967; Davidson, 1998; Magnan, 1986; Rifkin, 2003; and Thompson, 1996) has shown that students completing three or four years of foreign language study at the university level "typically demonstrate oral proficiency in the intermediate range" and therefore the IH level corresponds most closely to the classroom concerns for most teachers. Moreover, my interest in the IH level stems also from its proximity to the level at which learners attain minimal working competency in a foreign language, advanced low, according to federal employment standards for employees whose jobs require any degree of language expertise.

I analyze the lexical complexity of the discourse obtained in these two tests in order to understand the differences and similarities in speech produced in each test. That is, a high proportion of lexical complexity will provide evidence that the tests measure similar communicative samples and therefore may be used for substitution, whereas a low proportion will show the opposite. The results of my research will help language professionals better understand the difference between interpersonal and presentational speaking in the learning and teaching dynamic and in the area of testing and teaching Russian as a foreign language.

Data Collection and Criteria for Analysis

My approach to the analysis of lexical complexity of language samples was based on previous work by Shohamy (1994) and Halliday (1989). In this article the term “lexical complexity” will refer to a combination of lexical density and lexical diversity of speech. Lexical density measures the balance between “content” (lexical tokens) and “function” (grammatical tokens) words produced in each test, while lexical diversity measures the semantic variety of “content” and “function” words¹. Lexical diversity measurements are twofold. First, lexical diversity determines which test elicits a more diverse lexicon by means of calculating the number of similar versus diverse lexical and grammatical tokens produced by testees within the same type of test. Second, lexical diversity also establishes which test prompted a higher percentage of Less Frequent Lexical Items (LFLIs).

To perform the analysis of lexical diversity in this research I used Patrick Waddington’s *First Russian Vocabulary* (which consists of the 850 most commonly used words) as the lexical foundation for learners of Russian as a foreign language. In order to investigate lexical diversity, words used by testees in their responses will be classified as basic and non-basic. Basic words will be found in Waddington’s *First Russian Vocabulary*, while non-basic ones will not be included in his dictionary. Non-basic words are classified according to their part of speech (e.g., nouns, verbs, and adjectives). Henceforth, in this article, I will refer to words not listed in the Waddington lexicon as “less frequent lexical items” (LFLIs). A percentage of the higher number of LFLIs was calculated in relation to the aggregate number of words in each speech sample. In order to see what lexical developments take place at the IH level in the OPI and SOPI and to determine which mode, interpersonal or presentational, elicits a higher diversity of lexical items, first, all LFLIs were counted and lexical diversity (content versus function words) was established, then the words which were not included in Waddington’s dictionary were analyzed by categories (noun, verb, adjective, and adverb).

¹ “Content” words are nouns, full lexical verbs, adjectives, and adverbs. “Function” words are determiners (articles and quantifiers), prepositions, pronouns, numerals, conjunctions, interjections, existential ‘there’, the particle ‘to’, negative particles, auxiliary and modal verbs (Sityaev, 2000: 294).

Shohamy (1994) suggested that "language samples vary by the number of oral versus literate features they contain":

Literate language samples feature higher lexical density (i.e., the number of lexical items per clause), measured by the number of lexical items (i.e., content items rather than words) in relation to grammatical items (i.e., function items). Accordingly, texts, which are more literate, will contain larger numbers of lexical items as a higher level of sophistication, that is, per clause. The reverse will be true for texts of an oral nature, which rely more on grammar than on lexical items. Thus, the complexity of literate language is lexical, while that of oral language is grammatical. (109)

In her study, "The Validity of Direct Versus Semi-Direct Oral Tests" Shohamy (1994) explored the validity of two tests in Hebrew based on samples consisting of 10 OPIs and 10 SOPIs. One of the areas she examined was discourse features (a comparison of lexical versus grammatical density). Shohamy's comparison of language samples for content (lexical) versus function (grammatical) items per clause in the two tests showed that SOPI responses produced a higher lexical density, because they contained more nouns, verbs, adjectives, and adverbs than OPI speech, while OPI samples contained a broader range of grammatical items (prepositions, pronouns, numerals, conjunctions, interjections, negation, auxiliary and modal verbs). In her research the SOPI responses included 60% lexical and 40% grammatical items, whereas OPI responses featured the reverse: 40% lexical and 60% grammatical items. This ratio suggests the two modes elicit two different types of communication, oral in the OPI, and literate in the SOPI (according to Halliday's definition of oral vs. literate text).

I will compare the lexical complexity of students' speech in OPI and SOPI descriptions, examining most closely their production of content versus function items. My hypothesis, consistent with Shohamy's research, is that because the SOPI elicits presentational mode speech it will exhibit a higher production of content words versus function words than OPI descriptions. On the other hand, OPI samples, I would hypothesize, will have more function

words because the OPI simulates spontaneous oral conversation and this type of speech is characterized by greater grammatical complexity as opposed to lexical complexity. I will argue that the prompts used by both tests, the mode in which the language is elicited, and the students' psychological comfort affect their production of words during a speaking proficiency test.

The data for analysis consist only of fully formed lexical items; unfinished words (words that either did not take an ending or were abandoned in the middle of their production) did not enter the pool of data. In those cases when students slightly mispronounced words but the meaning of words was still recognizable in context, such words were included in the analysis. For example, if a testee intended to say *гостеприимный* *hospitable* but said *гостливый* then the word was counted towards the total number of words produced. However, if a testee produced lexical tokens that do not exist as words in Russian (*вешеле, рабатиками*) and the meaning of such tokens was not recognizable from the context, the tokens were eliminated from the data completely.

The data excluded categories that did not carry any semantic meaning such as fillers (including *там there, вот here*), fillers that are not words (for instance *а, э*), discourse markers (including *ну well*), interjections (for instance, *ой oh*), and words of dis/agreement in the beginning of responses (*да yes* and *нет no*).

As it was already mentioned earlier, there were three sources of data: ACTFL (OPI), CAL (SOPI) and Middlebury Russian Summer Program (both OPI and SOPI recordings). The OPI data from Middlebury consist exclusively of official OPIs conducted by certified testers. Middlebury SOPI data include interviews of those who volunteered at the Middlebury Russian School to participate in the research. First, I will analyze the entire pool of data obtained at the Intermediate High level. I will compare lexical density and lexical diversity of OPI vs. SOPI descriptions. After that, in order to confirm the findings with regard to which interview modality elicits a more lexically complex response, I will select and analyze OPIs and SOPIs of those Middlebury students who took both tests within the same week to avoid discrepancies in results due to rapid language acquisition in the context of the summer immersion program there. The description responses from these selected interviews will be referred to as responses of Middlebury group.

At the intermediate high level the entire pool of data consist of 18 audio-recorded speech samples (7 OPI and 11 SOPI), out of which 8 were from Middlebury group students at the IH level who agreed to take both the OPI and the SOPI (4 and 4).

Quantitative Data Analysis and Discussion

Tables 1, 2 and 3 show the results for the entire pool of data. Table 1 summarizes the total number of speakers and the total number of descriptions produced in each test at this level. Table 2 indicates the total number of tokens elicited by the OPI and the SOPI and the total number of LFLIs uttered by subjects in each test. Table 3 represents the same information as in Table 2 but per description per each test as a whole. The last four columns of Table 3 show the frequency of basic lexicon and LFLIs in each test in percentage.

Results gathered from the entire level show that in the SOPI, 11 subjects produced one solid instance of description per test, whereas 7 subjects in the OPI produced 32 instances of description. This is attributable to the fact that the OPI elicits description more than once based on the tester's assessment of the student's level. The mean number of descriptions per person in each test shows that, while in the SOPI subjects gave one description each, in the OPI each subject produced about 4-5 descriptions per interview. The aggregate number of descriptions at this level for the OPI is 3 times higher than the aggregate number of descriptions produced by subjects of the SOPI (32 vs. 11).

Table 1. Entire pool of data.

| | OPI | SOPI |
|--------------------|-----|------|
| Total speakers | 7 | 11 |
| Total descriptions | 32 | 11 |

The analysis of lexical density of speech at the IH level, in other words the balance between content (lexical) versus function (grammatical) items, showed that both tests elicited a similar ratio between content and function items at this level: 49% of content and 51% of function items in the OPI and nearly the reverse 52% vs. 48% in the SOPI.

Table 2. Quantitative results.**Total and unlisted OPI and SOPI items at the IH level. Entire pool of data.**

| IH | OPI | | SOPI | |
|--------------------------|-------------|------------|------------|------------|
| | No. words | No. LFLI | No. words | No. LFLI |
| Total words per level | 1731 | 153 | 858 | 133 |
| Number of content items | 851 | 141 | 448 | 124 |
| Number of function items | 880 | 12 | 410 | 9 |
| Content items | | | | |
| Number of nouns | 301 | 85 | 152 | 45 |
| Number of verbs | 198 | 19 | 96 | 10 |
| Number of adjectives | 118 | 27 | 108 | 60 |
| Number of adverbs | 234 | 10 | 92 | 9 |

Table 3. Quantitative results. Total and unlisted OPI and SOPI items and their frequency per description. Entire pool of data.

| IH | OPI | | SOPI | | OPI in % | | SOPI in % | |
|-----------------------------|--------------|-------------|-----------|--------------|--------------|-------------|--------------|-------------|
| | No. words | No. LFLI | No. words | No. LFLI | Fqcy. words | Fqcy. LFLI | Fqcy. words | Fqcy. LFLI |
| Total words per description | 54.09 | 4.78 | 78 | 12.09 | 100 | 8.83 | 100 | 15.5 |
| Number of content items | 26.59 | 4.4 | 40.72 | 11.27 | 49.16 | 8.14 | 52.21 | 14.45 |
| Number of function items | 27.5 | 0.38 | 37.27 | 0.81 | 50.84 | 0.69 | 47.79 | 1.04 |
| Content items | | | | | | | | |
| Number of nouns | 9.4 | 2.65 | 13.81 | 4.09 | 17.38 | 4.91 | 17.71 | 5.25 |
| Number of verbs | 6.18 | 0.59 | 8.72 | 0.9 | 11.43 | 1.09 | 11.18 | 1.16 |
| Number of adjectives | 3.68 | 0.84 | 9.81 | 5.45 | 6.81 | 1.55 | 12.58 | 6.99 |
| Number of adverbs | 7.31 | 0.31 | 8.36 | 0.81 | 13.51 | 0.57 | 10.72 | 1.04 |

The analysis of LFLIs at the IH level showed that out of 1731 total words produced by subjects in the OPI, 153 of them were of LFLIs, not included in the basic vocabulary list by Waddington. In the SOPI there were a total of 858 words of which 133 words LFLIs. In the OPI only 8.83 % of the description discourse featured LFLIs, while in the SOPI 15.5% of the total data was LFLIs, which is almost two times higher (see Table 3, columns 6 and 8).

Though both tests elicited a similar ratio of content vs. function items total at this level, the close analysis of the lexical density of LFLIs shows that in both tests testees produced content LFLIs 11 times more often than function LFLIs (8.14% to 0.69% in the OPI, and 14.45% to 1.04% in the SOPI). In other words, in the OPI only 8% of the entire number of content items was LFLIs, while in the SOPI more than 13% were LFLIs.

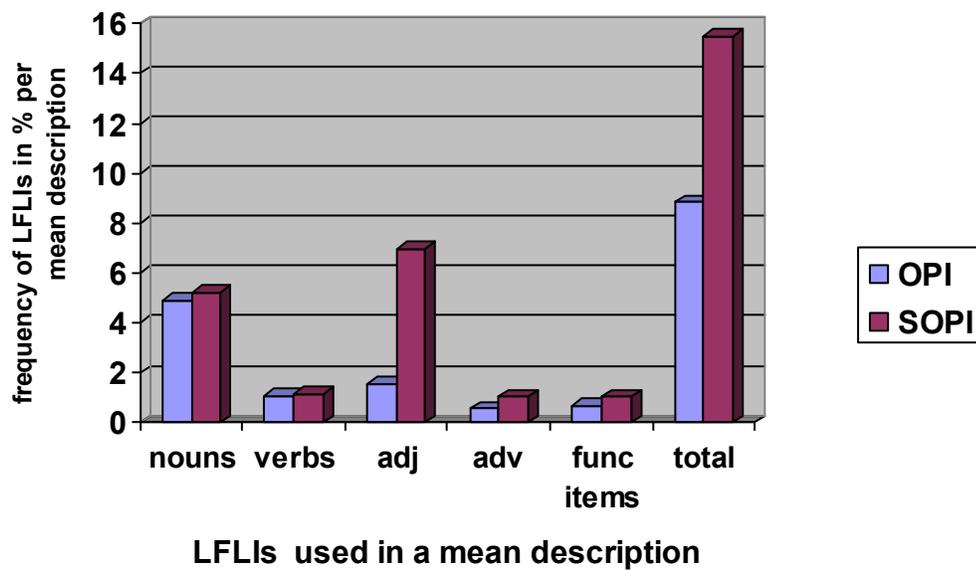
The analysis of the entire data lexical diversity of content words at this level showed a high ratio between the nouns and verbs (5th and 7th column in Table 3, also Figure 1): about 17.5% per nouns (17.38% in the OPI and 17.71% in the SOPI) and about 11.10% per verbs (11.43% in the OPI and 11.18% in the SOPI). Meanwhile, the ratio between adjectives and adverbs in the OPI and the SOPI was low. There were two times more adjectives used in the mean SOPI description than in the mean OPI description: 6.81 to 12.58. The frequency of adverbs in these tests was almost the same: 13.51 in the OPI to 10.72 in the SOPI.

The analysis of the entire data lexical diversity of content items only within the LFLIs showed different results (Table 3, column 6 and 8): though there was a high ratio within the OPI and the SOPI with regard to elicitation of nouns and verbs (4.91 to 1.09 in the OPI and 5.25 to 1.16 in the SOPI), and the ratio among adjectives, adverbs and function items was low. The visual results of the lexical diversity of LFLIs are given in Figure 1. The data suggest that the SOPI elicited almost twice as many adverbial and function LFLIs and almost four times more adjectival LFLIs than the OPI. Compare the OPI elicited 1.55% of adjective LFLIs per description while the SOPI elicited almost four times more, 6.99%.

Based on the quantitative results of basic words in the entire pool of data, the initial conclusions of this research are: first, the analysis above suggests that both tests elicit lexical density (the ratio between content and function items) with near identical frequency, about 50% of content and 50% of

function items in each test, suggesting that the two tests are comparable and both modes elicit similar kinds of discourse in terms of the speakers' lexicon. However, the close analysis of entire LFLIs data showed that the SOPI elicited more diverse LFLIs: at nearly double the frequency of the OPI (8.83% to 15.5%). Second, the comparison of lexical diversity within only LFLIs provided evidence that both tests elicited nouns and verbs with similar frequency but the SOPI did a better job of eliciting significantly more lexically diverse adjectival, adverbial and function LFLIs. This may ultimately suggest that the SOPI is better at eliciting lexical diversity of speech than the OPI but there are some extenuating circumstances that need to be considered. I will turn now to the qualitative, type-token, analysis of the results obtained for lexical density of LFLIs in the OPI and the SOPI.

Figure 1. Frequency of LFLIs content (nouns, verbs, adjectives and adverbs) and function items used by IH level speakers in the OPI and the SOPI descriptions. Entire pool of data.



Type-Token Data Analysis and Discussion

An explanation for the previous results may lie in the relationship between the language samples and the prompts that elicit them. While the OPI prompts ask testees to describe or compare places, objects and people, vocabulary used in the prompts cannot serve as an input or source vocabulary for the response and

does not obligate students to use particular lexicon in their discourse. Examples of OPI prompts include: Расскажи мне об Алабаме. Я там никогда не была. *Tell me about Alabama. I have never been there.* Расскажите от этой квартире. Опишите её. *Tell me about this apartment. Describe it for me.* Опишите мне брата, пожалуйста. *Describe your brother, please.* Расскажите мне о Вашем отце. Вы сказали, он художник. Какой у него характер? *Tell me about your father. You said he is an artist. Tell me about his personality.*

The SOPI prompt on the other hand asks testees to compare the advantages and disadvantages of studying at a private University or College: Как ты думаешь, в чём преимущества и недостатки обучения в частном университете? The question elicited not only the vocabulary chosen by the testees but also the words used in the prompt itself. Thus, the prompt in the SOPI to some extent cued testees to use the vocabulary of the prompt including words such as 'public' vs. 'private', and 'advantages' and 'disadvantages', some of which (like 'private', 'advantages' and 'disadvantages') were mentioned in the prompt itself. This has a direct effect on the results because these words are not included in Waddington's dictionary whereas many of the words likely to be found in a description of one's apartment are. Thus, the structure of the SOPI allowed testees to incorporate these lexical items into their speech after they heard the prompt first in English and then in Russian. Testees taking the SOPI had a slight advantage over those taking the OPI because if SOPI testees did not know Russian words for 'private', 'advantages' and 'disadvantages' used in the input/ prompt or could not recall these lexical items in Russian prior to their response, testees still were able to hear the core vocabulary, necessary to handle the situation successfully in Russian and were able to intake these lexical items and use them in their output. Since the prompt focused on the advantages and disadvantages of studying at a private or state university, consequently, the lexical items 'private', 'state/ public', 'advantages' and 'disadvantages' were frequently used by testees in the comparison of the two types of institutions.

Another issue affecting the results is the type of token (in other words, the repetition of certain words) each counted as a single instance of LFLIs including the words mentioned above that are present in the prompt. Table 3 represents the type-token results of the lexical diversity of LFLIs.

The type-token analysis of lexical diversity of LFLIs showed that out of 8.83% of LFLIs in the OPI, 6.58% (or three quarters) were diverse and 2.25%

(one quarter) of the vocabulary were repeated. In the SOPI out of 15.5% of LFLIs, half (7.45%) were repeated. In the SOPI the major repetition of lexical items took place in the categories of verbs, adjectives, adverbs and function items. (Compare the frequency of words in the SOPI: total elicited LFLIs vs. their diversity, two right columns in table 4).

The higher results suggest that in the SOPI the prompt provided testees with the lexical items necessary to handle the task. The SOPI prompt predetermined a high frequency of utilization of the lexical items used in the prompt to be used in testees' responses. Now compare the results of lexical diversity in the two tests (table 4). In the SOPI, out of 45 nouns only 27 were different, in other words nearly half of the nouns accounting for the high number were repeated: for example образование *education* was repeated 6 times, преимущество *advantage* was repeated 5 times. Out of 10 LFLIs verbs used in the SOPI, only half of them were diverse (for example, the verb поступить *to enter a university* alone was used 4 times among 11 subjects). So the previous contention that the SOPI is better in eliciting lexical diversity is called into a question. Consider Figure 2 which shows the comparison of frequencies of lexical density with repeated items included and then with them factored out. When the repetition is factored out such that only unique content (lexical) items are counted, the picture of the lexical diversity of LFLIs changes significantly.

Table 4. Type-token results, Lexical Diversity of LFLIs in the OPI and SOPI lexicon at the IH level. Entire pool of data.

| | | OPI No. LFLI | | SOPI No. LFLI | | OPI Fqcy. LFLI in % | | SOPI Fqcy. LFLI in % | |
|-------------------------|------------|--------------|------------|---------------|-----------|---------------------|-------------|----------------------|-------------|
| | | total | diverse | total | diverse | total | diverse | total | diverse |
| Content items | Nouns | 85 | 62 | 45 | 27 | 4.91 | 3.58 | 5.25 | 3.14 |
| | Verbs | 19 | 18 | 10 | 5 | 1.09 | 1.03 | 1.16 | 0.58 |
| | Adjectives | 27 | 22 | 60 | 11 | 1.55 | 1.27 | 6.99 | 1.28 |
| | Adverbs | 10 | 8 | 9 | 8 | 0.57 | 0.46 | 1.04 | 0.51 |
| Function items | | | | | | | | | |
| | | 12 | 4 | 9 | 3 | 0.69 | 0.22 | 1.04 | 0.34 |
| total LFLI words | | 153 | 114 | 133 | 64 | 8.83 | 6.58 | 15.5 | 7.45 |

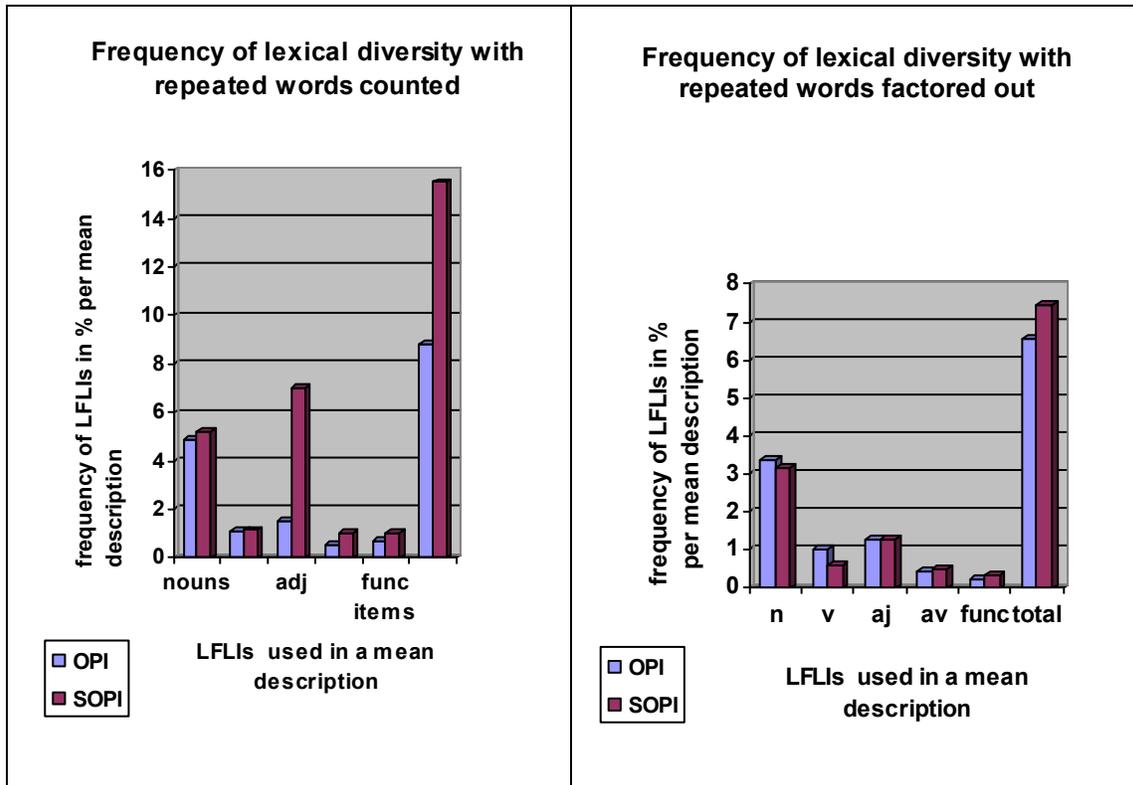
The analysis of adjectives exhibited striking results. Out of 60 LFLIs adjectives used in the SOPI, half of them were частный *private* (33 usages), one fourth were государственный *public* (15 usages) and 5 usages were публичный also intended to mean 'public', actually a semantic misuse of the item². Thus, the SOPI test of presentational mode speech elicited only 11 semantically various adjectives.

The type-token analysis of lexical diversity of LFLIs in the OPI showed higher results. For example, out of 27 adjectives only 5 were repeated (2 строгий *strict*, 2 безумный *crazy* and 3 худой *skinny*), in other words 22 adjectives in the OPI LFLIs were semantically diverse vs. only 11 semantically diverse adjectives in the SOPI LFLIs. Out of 85 nouns only about 10 of them were used twice, and only one verb was repeated twice вырос *I grew up*. Figure 2 below shows the comparative results of the frequency of lexical diversity of content and function LFLIs used by IH level speakers in the OPI and the SOPI descriptions. Left figure represents the frequency of lexical diversity with repeated words counted, while the right figure shows the frequency of lexical diversity with repeated words factored out.

Despite the fact that the SOPI prompt elicited more (almost 2 times more) of LFLIs lexical density than the OPI (15.5% to 8.8%), a close look at the lexical diversity of LFLIs in description (type-token analysis) showed that both tests elicited similar lexical diversity of speech in the LFLIs at the IH level (6.58% in the OPI and 7.45% in the SOPI). This shows that while the SOPI testees repeated the same vocabulary, much of it provided by the prompt, the OPI testees used more varied original lexicon. The quantitative analysis (the analysis of the total number of tokens) shows that the frequency of lexical diversity in the SOPI was higher among nouns, verbs and particularly adjectives, while the type-token analysis of speech showed that the OPI elicited slightly higher lexical diversity of nouns (3.58 to 3.14), similar lexical diversity of adjectives and two times more diverse verbs: 1.03 vs. 0.58!

² The correct adjective is государственный, and the incorrect usage of публичный *public/state* proves my earlier hypothesis that in the SOPI students hear the prompt both in English and Russian, and the prompt provides students with the necessary vocabulary. The word 'state/public' was not in the prompt, therefore students had to come up with the Russian equivalent, and in some cases students who did not know the correct equivalent to the word 'state' or 'public' used a false cognate 'public' публичный, which in Russian carries a different meaning.

Figure 2. Comparative results of the frequency of lexical diversity of content and function LFLIs used by IH level speakers in the OPI and the SOPI descriptions. Entire pool of data.



Duration, the time available for testees to respond to a prompt, was another factor influencing the outcomes. Despite the relative freedom of time allowed to OPI testees per response (description), the mean number of lexical items produced per description in the OPI was almost 20 items less than in the SOPI: 54 to 78 (see table 5). A larger number of lexical items per response in the SOPI could be attributable to the presentational mode in which testees perform. In the SOPI, first, testees hear the situation to which they need to give a response, and then the instructions allow them 20 seconds to plan their responses. After the planning time is over, testees hear the prompt in the target language and then they are given a fixed amount of time, differing by prompt, to give their response. The interpersonal modality of the OPI excludes the reflection and preparation time. Second, the SOPI format has a period of 1 minute and 20 seconds (for the task of description) during which students may feel they need to “fill the silence” with further description. This might account for much of the repetitiveness of the response, when a testee has nothing

further to say but feels he or she must keep talking. Regardless of level, all SOPI testees are given the same amount of time to perform the task of description. In other words, it is assumed that in order to show that a speaker can perform Advanced level functions; he/she is provided with a certain amount of time for that. Nevertheless, since IH speakers perform advanced-level tasks at least 50% of the time, but not consistently, they typically give shorter responses than Advanced level speakers. Thus, in the SOPI, when IH level speakers complete their answer to the prompt before the time for the answer expires, they hear silence on the tape. Some speakers may feel compelled to produce more speech, and longer responses are sometimes the result. In the OPI, testers do not have a set time during which testees may plan and then give a response to a particular prompt. The flow of the OPI is that of a natural conversation, in which one person (a tester) asks questions and the other (a testee) gives a response. The testee is not given time to plan and organize his/her discourse. In this setting, testees feel pressure to give an immediate response to avoid a pause between the question and their response.

When looking at the rate of speech, the number of words per minute, the opposite is true. Subjects produced a higher rate of speech, in other words more lexical items per minute in the interpersonal mode than in the presentational. The mean OPI response was 41.5 seconds long while the mean SOPI response was 68 seconds long. Thus, per minute OPI speakers produced a mean number of 78 words vs. almost 69 words per minute in the SOPI (see table 5).

Table 5. Rate of speech in seconds (mean length of description and mean number of words per minute) at the IH level. Entire pool of data.

| IH level | OPI | SOPI |
|---------------------------------------|-------|-------|
| Mean length of description in seconds | 41.5 | 68.00 |
| Mean number of words per description | 54.09 | 78.00 |
| Mean number of words per minute | 78.00 | 68.82 |

Such results are not surprising as the interpersonal mode presented subjects with an opportunity to negotiate meaning and seek paralinguistic assistance. In other words, the presence of the interlocutor affected the language production. Subjects were able to see by the reaction of the tester whether the information presented in the response was sufficient or needed extension. It was noticed that encouragements and phrases like “really”, “how

interesting”, “right”, inserted by OPI testers during testees’ responses prompted OPI speakers to produce more language. The presentational mode of the SOPI excluded this advantage. Though the presence of the “live” interaction featured a higher number of lexical items per minute, the interpersonal mode did not have any impact on the higher frequency of LFLIs in learner discourse.

Table 6. Number of speakers and number of descriptions produced in the OPI and the SOPI. Middlebury group.

| | OPI | SOPI |
|--------------------|-----|------|
| Total speakers | 4 | 4 |
| Total descriptions | 17 | 4 |

Unique findings were obtained from the Middlebury group. First, the number of tokens and the type-tokens comparison of lexical density and lexical diversity across tests in the Middlebury group showed slightly different results from those obtained for the entire level. Meanwhile in the Middlebury group both tests elicited higher lexical density of function items, the ratio of function items in the OPI was significantly higher than the ratio in the SOPI (there were 62.72% of function items in the OPI and only 55.70% in the SOPI). This finding may suggest that the nature of the OPI is more orate while the nature of the SOPI is more literate. Second, similar to the results of the entire level, the SOPI elicited higher frequency of LFLIs than the OPI (9.63% in the OPI to 12.99% in the SOPI) in the Middlebury group. However a detailed analysis of LFLIs among Middlebury group and the entire pool of data showed a slight difference. In the Middlebury group the OPI mode elicited much higher frequency of verbs while the frequency of nouns and adjectives is similar to the frequency results obtained for the entire level: almost exact frequency of nouns and almost two times more adjectives in the SOPI than in the OPI mean description. Table 6, 7 and 8, and Figure 3 show the results for the Middlebury group.

The close analysis of lexical diversity of LFLIs in the Middlebury group supports the findings for the entire level: the SOPI elicits almost two times more adjectival and adverbial LFLIs but slightly less noun and verbal LFLIs than the OPI. Table 9, columns number 6 and 8, and figure 5 demonstrate comparative results of the frequency of lexical diversity of content and function LFLIs used

by IH level speakers in the Middlebury group in the OPI and the SOPI descriptions.

Table 7. Quantitative results. Total and unlisted OPI and SOPI items at the IH level. Middlebury group.

| IH | OPI | | SOPI | |
|--------------------------|-------------|------------|------------|-----------|
| | No. words | No. LFLI | No. words | No. LFLI |
| Total words per level | 1038 | 100 | 377 | 49 |
| Number of content items | 387 | 92 | 167 | 44 |
| Number of function items | 651 | 8 | 210 | 5 |
| Content items | | | | |
| Number of nouns | 173 | 54 | 63 | 15 |
| Number of verbs | 141 | 16 | 32 | 4 |
| Number of adjectives | 48 | 18 | 43 | 22 |
| Number of adverbs | 25 | 4 | 9 | 3 |

Table 8. Quantitative results. Total and unlisted OPI and SOPI items per description. Middlebury group.

| IH | OPI | | SOPI | | OPI in % | | SOPI in % | |
|-----------------------------|--------------|-------------|--------------|--------------|--------------|-------------|--------------|--------------|
| | No. words | No. LFLI | No. words | No. LFLI | Fqcy. words | Fqcy. LFLI | Fqcy. words | Fqcy. LFLI |
| Total words per description | 61.05 | 5.88 | 94.25 | 12.25 | 100 | 9.63 | 100 | 12.99 |
| Number of content items | 22.76 | 5.41 | 41.75 | 11 | 37.28 | 8.86 | 44.30 | 11.67 |
| Number of function items | 38.29 | 0.47 | 52.5 | 1.25 | 62.72 | 0.77 | 55.70 | 1.33 |
| Content items | | | | | | | | |
| Number of nouns | 10.17 | 3.17 | 15.75 | 3.75 | 16.67 | 5.20 | 16.71 | 3.98 |
| Number of verbs | 8.29 | 0.94 | 8 | 1 | 13.58 | 1.54 | 8.49 | 1.06 |
| Number of adjectives | 2.82 | 1.05 | 10.75 | 5.5 | 4.62 | 1.73 | 11.40 | 5.84 |
| Number of adverbs | 1.47 | 0.23 | 2.25 | 0.75 | 2.40 | 0.39 | 2.39 | 0.80 |

Figure 3. Frequency of LFLIs content (nouns, verbs, adjectives and adverbs) and function items used by IH level speakers in the OPI and the SOPI descriptions.

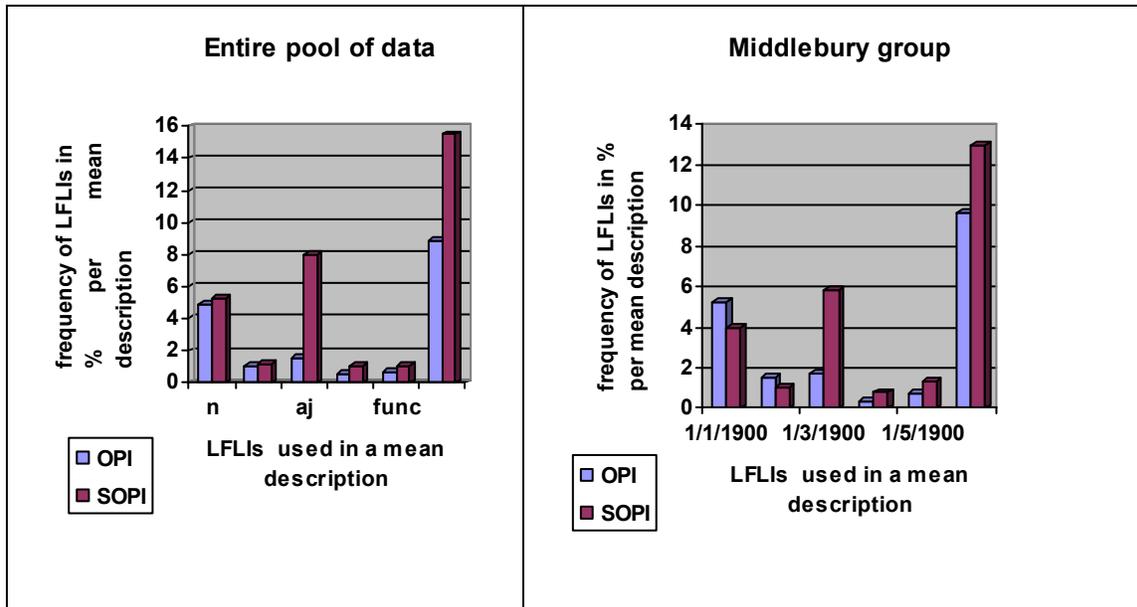
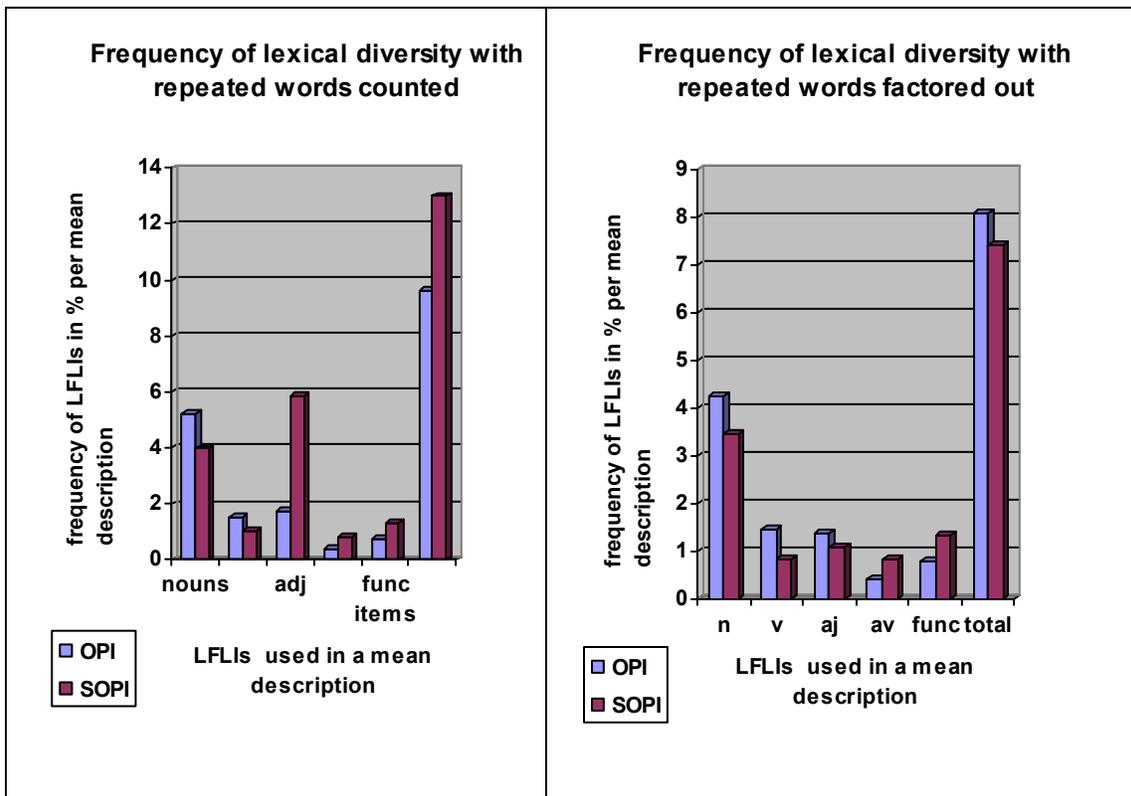


Table 9. Type-token results, Lexical Diversity of LFLIs in the OPI and SOPI lexicon at the IH level. Middlebury group.

| | | OPI No. LFLI in group | | SOPI No. LFLI in group | | OPI Fqcy. LFLI in % in group | | SOPI Fqcy. LFLI in % in group | |
|-------------------------|------------|-----------------------|-----------|------------------------|-----------|------------------------------|-------------|-------------------------------|-------------|
| | | total | diverse | total | diverse | total | diverse | total | diverse |
| Content items | Nouns | 54 | 44 | 15 | 13 | 5.20 | 4.24 | 3.98 | 3.45 |
| | Verbs | 16 | 15 | 4 | 3 | 1.54 | 1.45 | 1.06 | 0.80 |
| | Adjectives | 18 | 14 | 22 | 4 | 1.73 | 1.35 | 5.84 | 1.06 |
| | Adverbs | 4 | 4 | 3 | 3 | 0.39 | 0.39 | 0.80 | 0/80 |
| Function items | | 8 | 7 | 5 | 5 | 0.77 | 0.77 | 1.33 | 1.33 |
| total LFLI words | | 100 | 84 | 49 | 28 | 9.63 | 8.09 | 12.99 | 7.43 |

Figure 5. Comparative results of the frequency of lexical density and lexical diversity of content and function LFLIs used by IH level speakers in the OPI and the SOPI descriptions. Middlebury group.



Conclusions

The results of this case study suggest that despite the larger number of descriptions and words produced at the IH level in the OPI, the SOPI descriptions included almost twice as many LFLIs, words not on Waddington's basic list, and a higher frequency of words per response. The explanations for this difference lie, I would argue, in the modality of the test and the prompt. The SOPI is presentational, and therefore provides testees time to prepare an answer. The OPI is interpersonal, allows for negotiation of meaning, lacks a preparation period, and is conducted exclusively in the target language.

However, because of the presence of certain lexical items in the SOPI prompt, the usage of which was necessary to perform the task, the higher frequency of lexical items in the SOPI seems to have been artificially inflated enabling SOPI speakers to out-perform OPI speakers by the frequency of lexical

density of LFLIs (Table 4). The task used in the SOPI created expectations for testees to use this vocabulary in their responses, providing them with English translations in the written materials accompanying the oral prompts. If students did not know or could not remember the core words required to complete the task in the SOPI, they were still able to give a response, using the LFLIs.

The comparative results of content (nouns, verbs, adjectives and adverbs) versus function (prepositions, pronouns, numerals, negation, conjunctions, interjections, and modal verbs) items of the present study differ from those found by Shohamy. The data from the Middlebury group suggest that both interpersonal and presentational modes elicited significantly more function than content items (62.72 in the OPI and 55.70 in the SOPI), though the entire pool of data showed opposite results suggesting that both modes elicit similar type of lexical complexity: half content items and half function items. This leads me to reject my original hypothesis and puts into question the findings of Shohamy in her earlier study comparing the OPI and SOPI in Hebrew. My results of the entire pool of data suggest that learner discourse was not affected by the mode of the test but rather by the prompt used in it.

The present study, of course, has limitations. First, the findings might have been different had the analysis of less frequent lexical items been based not on Waddington's *First Russian Vocabulary* but on another dictionary. For example, Waddington's dictionary lists terms which are out of date in contemporary Russian culture and which have low frequency in today's daily usage, such as 'pioneer', 'comrade', but does not list contemporary words cognates like 'computer', 'technology, equipment', 'tourist', 'politics' and etc. The purpose of this research was to obtain a rough estimate of the overall lexical complexity and this procedure was accomplished. In future research, lexical complexity of testees' discourse might be based on a more up-to-date dictionary, such as Brown's *Russian Learner's Dictionary: 10,000 Words in Frequency Order*, and on the vocabulary provided in the most commonly used college-level textbooks for beginning Russian such as *Golosa, Nachalo, Russian Stage One: Live from Moscow*, and *Troika*.

Second, the current research focused on LFLIs exclusively. In the future it might be extended to the analysis of both basic and non-basic vocabulary in these tests.

Third, the OPI measures linguistic knowledge, grammar and sentence structures of the language (Halleck, 1995), but it does not assess language complexity. The OPI does not have a mechanism to diagnose the lexical complexity of responses as the ACTFL OPG have no lexicon criterion. My observations in this area are unattached to the guidelines and the test criterion; nonetheless they remain interesting and suggest further investigation in other languages.

Finally, all of this research has been conducted with a small set of speech samples. To be certain of the validity of these findings, one would need a larger set of official data: at present this is not possible in Russian SOPI since instructors who purchase the SOPI and the Rater Training kits can administer the test to any number of students without reporting the ratings to CAL. (from a personal correspondence with Meg Malone, February 2007). Perhaps the day will come when scholars can pick up this preliminary study and validate or refute its findings.

Bibliography

Brecht, Richard; Davidson, Dan, & Ginsberg, Ralph. *Predictors of Foreign Language Gain During Study Aboard*. Washington, DC: National Foreign Language Center Occasional Papers (1993).

Carroll, John. "Foreign Language Proficiency Levels Attained by Language Majors Near Graduation from College". *Foreign Language Annals* 1 (1967): 131-151.

Davidson, Dan. "Summer, semester, and academic year proficiency profiles". Paper presented at the *Annual AATSEEL Conference*, San Francisco, CA (1998).

Halleck, Gene. "Interrater reliability of the OPI. Using Academic Trainee Raters". *Foreign Language Annals* 29, no 2 (1995): 223-238.

Halliday, M.A.K. *Spoken and Written Language*. Oxford: Oxford University Press, 1989.

- Magnan, Sally. "Accessing Speaking Proficiency in the Undergraduate Curriculum: Data from French". *Foreign Language Annals* 19 (1986): 429-438.
- Center for Applied Linguistics. *Rater training kit: Russian speaking test, 1st ed.* Washington, DC, 1998.
- Rifkin, Benjamin. "Oral Proficiency Learning Outcomes and Curricular Design". *Foreign Language Annals* 36, no 4 (2003): 582-588.
- Schiffrin, Deborah. *Discourse markers*. Cambridge, UK: Cambridge University press, 1987.
- Shohamy, Elana. "The validity of direct versus semi-direct oral tests". *Language Testing* 11, no 2 (1994).
- Sityaev, Dmitry. "The relationship between accentuation and information status of discourse referents: A corpus-based study". *UCL Working Papers in Linguistics* 12, London. (2000).
- Swender, Elvira. *ACTFL Oral Proficiency Interview tester training manual*. Yonkers, NY: American Council on the Teaching of Foreign Languages, 1999.
- Thompson, Irene. "Assessing Foreign Language Skills. Data from Russian". *Modern Language Journal* 80, (1996): 47-65
- Waddington, Patrick. *First Russian Vocabulary*. Duckworth Publishing, 2003.