Semantic paraphasias in normal speech: A corpus-based analysis

Svetlana Gorokhova
St Petersburg State University

Abstract

Semantic paraphasias naturally occurring in Russian normal speech were analysed in terms of word frequency, word length, word association norms, and cooccurrence strength. Frequencies of target words were found to be significantly lower than frequencies of their substitutes; besides, there is a very significant positive correlation between target and error frequency values. Contrary to the view that the frequency effect is located at the stage of phonological encoding, the results suggest that frequency is coded at the earlier stage of lexical selection. Target words appear to have different frequency and length characteristics depending on whether they elicit taxonomically or non-taxonomically related errors. Word length is a significant variable that determines the outcome of the error for non-taxonomically related target-error pairs but not for taxonomically related pairs. At the same time, taxonomically related target-error pairs are characterized by much higher co-occurrence measures and stronger associative links compared to non-taxonomically related pairs. Theoretical implications of these findings are discussed.

Introduction

The available data on the role and the locus of the frequency effect in lexical retrieval are controversial. While many authors argue that the frequency effect is located at the stage of accessing phonological forms (Jescheniak & Levelt 1994; Jurafsky 2003 etc.) rather than the semantic lemma level, there is some evidence to suggest that lexical selection (lemma retrieval) is also affected by word frequency (Finocchiaro and Caramazza 2006; Navarrete, Basagni, Alario, & Costa 2006).

One of the ways to investigate the role and locus of the frequency effect in language production is to compare the frequencies of the target and substitute words in semantic paraphasias that occur in normal and aphasic speech. Hotopf (1980)'s study failed to discover any significant frequency difference between targets and substitutes in semantic substitution errors, leading him to conclude that both semantically related words involved in a substitution belong to the same frequency range. Likewise, del Viso, Igoa, & Garcia-Albea (1991) did not find any target-error frequency difference in semantic substitutions in their study of Spanish speech errors. Similarly, Harley & MacAndrew (2001) found that neither semantic nor phonological targets are replaced by more frequent intrusions in English speech errors. They argue that word frequency mainly affects phonological errors. In a study of a dyslexic patient's speech, Gerhand & Barry (2000) found no significant frequency effect. Conversely, German & Newman (2004) found that children with word-finding difficulties were more likely to produce substitutions that were higher in frequency of occurrence than the target word. Gordon (2002) concludes that while item frequency influences a word's susceptibility to error, its effect on error outcome is questionable, at least not in normal speech.

In Kittredge, Dell, Verkuilen, & Schwartz (2008)'s analysis of semantic word errors made by aphasic patients, an increase in the log-transformed frequency of the target significantly predicted an increase in error log frequency. They interpreted this correlation as being indicative of a competitive, frequency sensitive lexical access process and concluded that both abstract and phonological representations of higher-frequency words may be more easily accessed than those of lower frequency words. These results have found some supportive evidence in Karatsinski (2010)'s study of word interruptions in production: while not concentrating on errors of performance, the author argues that "frequent words are easier to access, faster to produce, and harder to interrupt than rare words".

Methods

This paper presents a study-in-progress report of a corpus analysis of about 300 Russian meaning-based noun substitution speech errors (slips of the tongue), i.e. cases when a semantically related error noun is substituted for the target noun. The errors were collected by tape-recording and digitally recording everyday conversations, telephone conversations, and live TV and radio programs such as talk shows and interviews.

To explore the role of frequency effects in lexical selection, the lexeme (token) frequency (i.e. the frequency of the particular inflected word form involved in the substitution) and the lemma frequency (i.e. the total frequency of all inflected forms) of each target noun were compared to those of the substitute (error) noun. The raw frequency value of a word was its frequency of occurrence in the Russian National Corpus (http://ruscorpora.ru). The results were analysed separately for the two groups of semantic substitutions, taxonomically related (category coordinate) and non-taxonomically related (non-category coordinate) target-error pairs (duplicate examples, i.e. examples of the same target-error lemma or lexeme pairs, were removed from the sets of examples). To study the effect of word length (measured in syllables) on semantic substitutions, target lexeme lengths were compared to error lexeme lengths.

It is generally assumed that log frequency is better correlated with many measures of performance than raw frequency (Oldfield & Wingfield 1965; Vitevitch 1997; see Kittredge et al. 2008 etc.); besides, log transformation is a usual way to normalise a non-Gaussian distribution, so in most tests, log transformed frequency values were used for comparisons (where log transformation was not applicable due to a big number of zero frequency values, square root transformation was used), but in the section that reports a comparison of target frequencies to error frequencies, the results for both raw and log transformed frequency values are included. For word length comparisons, log-transformed word lengths were used. Target-error co-occurrence measures were computed basing on raw frequency values as per the formulas.

Data

A number of experimental studies have indicated a special status of category coordinate relations between words in picture naming tasks (e.g. Kroll & Stewart 1994; Damian & Als 2005). To find out whether category coordinate relations between the target and the substitute have an impact on frequency characteristics of semantic substitution errors, the cases where the target word and its substitute are taxonomically related (category coordinates) were analysed separately from the examples that represent other types of semantic relatedness between target and error words. The error corpus under study comprised 177 taxonomically related (category coordinate) and 114 non-taxonomically related target-error pairs.

Below are some representative examples of taxonomically related and non-taxonomically related target-error pairs.

Taxonomically related

(a) Kak ty dumaeš′, esli ja etot tort ukrašu dol′kami **apel′sina**? → …**limona**?
What do you think – if I decorate the cake with segments of an **orange** → …**lemon**? (*fruit*)

(b) Kakie krasivye **gvozdiki**! → …**rozočki**!
What beautiful **carnations**! → **roses**! (*flowers*)

(c) …Deržit **gitaru** tak, kak budto eto violončel′ → … Deržit **flejtu**…
…holds the **guitar** as if it were a cello → …holds the **flute…** (*musical instruments*)

(d) Tri **goda**, tri **goda** devočke → tri **časa**…
The girl is only three **years** old, just three → …three **hours**… (*units of time*)

(e) Ja tebe **bljudce**, meždu pročim, xoču dostat′ → … **tarelku** …
Incidentally, I want to get you a **saucer** → … a **plate (***tableware*)

Non-taxonomically related

(a) Koška vse vremja deret **kovry** → …**poly**
The cat keeps tearing the **carpets** → …the **floors**

(b) A eto pal′to ona ešče **zimoj** budet nosit′ → …**nočju**…
She will wear this coat **in winter**, too → …**at night**…

(c) Nakanune **pobedy** tože pogibali naši soldaty → Nakanune **vojny**…
Our soldiers were dying on the eve of the **victory**, too → …on the eve of the **war**…

(d) Ja podhožu, a ona ležit veselen′kaja i **rešetku** gryzet → …**setku**…
I came up, and there she was lying, looking cheerful and gnawing at the **bars** →…at the **netting**

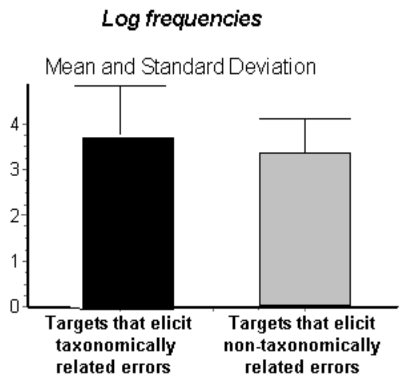
(e) Rasprodaža v **janvare** → …v **ponedel′nik**
The sale is **in January** → …**on Monday**

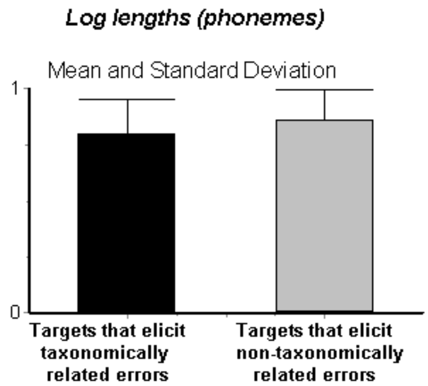
Results

1. On average, targets that elicit taxonomically related errors have higher frequencies and shorter lengths than targets that elicit non-taxonomically related errors.

An unpaired t-test was run to compare the log-transformed lemma frequencies of target words that elicit taxonomically related errors to the lemma frequencies of those targets that elicited non-taxonomically related errors. Since the SDs of the two groups of nouns are significantly different (SD = 0.9793 for targets that elicit taxonomically related errors and SD = 0.7409 for targets that elicit non-taxonomically related errors), Welch correction was applied to the unpaired t-test. The test revealed an extremely significant frequency difference between the two groups: t (232) = 3.144 (p = 0.0019). The mean log-transformed lemma frequency of the *non-taxonomically related* group targets is 3.412 (n = 100) while the corresponding mean value for the *taxonomically related* group targets is 3.766 (n = 135).

A nonparametric Mann-Whitney test run to compare the log-transformed lexeme lengths (measured in phonemes) of the targets from the taxonomically related group to those of the targets from the non-taxonomically related group (an unpaired t-test was not appropriate because the distribution of log length values did not pass a normality test) revealed a very significant difference (Mann-Whitney U-statistic = 6428.5, p = 0.0015, mean target log length within *taxonomically related* group = 0.7997, mean target log length within *non-taxonomically related* group = 0.8633).

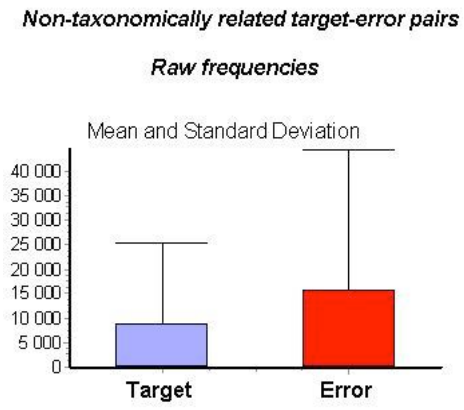


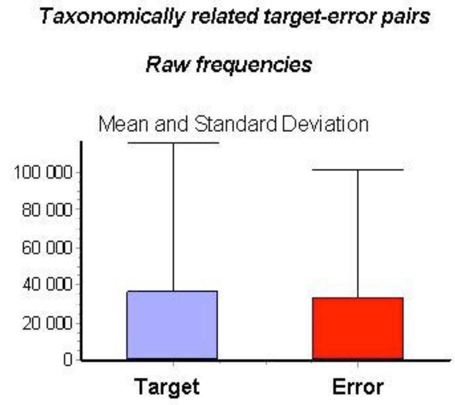


2. Unlike other studies that failed to find a significant frequency difference between target and error frequencies of semantic paraphasias, Russian speech error evidence indicates that **higher-frequency nouns tend to substitute for lower-frequency semantically related nouns.**

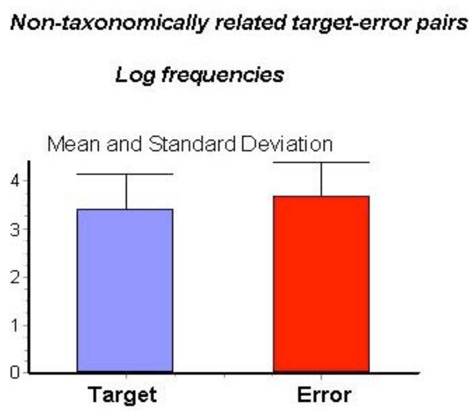
Since the **raw frequency values** do not appear to follow a Gaussian distribution, a Wilcoxon matched-pairs test was used to compare target and error frequencies. Within *overall* semantic substitution errors, the target-error raw frequency difference is very significant (for lemma frequencies, W (sum of ranks) = –6205.0, p = 0.0033, mean target frequency = 24895, mean error frequency = 25901; for lexeme frequencies, W = –8309.0, p = 0.0008, mean target frequency = 6592.2, mean error frequency = 6971.4). This result, however, must be due to the very significant difference between the target and error lemma raw frequencies within the *non-taxonomically related* group (for lemma frequencies, W = –1808.0, p = 0.0019, mean target frequency = 8868.5, mean error frequency = 15595; for lexeme frequencies, W = –1998.0, p = 0.001, mean target frequency = 1657.1, mean error frequency = 3241.2).

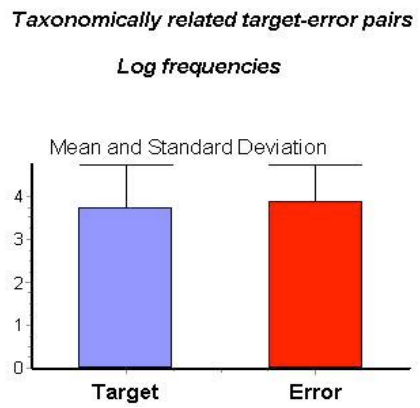
A comparison of the target and error raw frequencies within the *taxonomically related* group yields a completely different result: it does not reveal any significant difference between target and error frequencies. Conversely, both for lemma and for lexeme frequencies, the mean raw target frequency appears to exceed the mean error frequency (for lemma frequencies, W = –994.00, p = 0.2914, mean target frequency = 36593, mean error frequency = 32957; for lexeme frequencies, W = –2124.0, p = 0.0732, mean target frequency = 9710.7, mean error frequency = 9328.5).





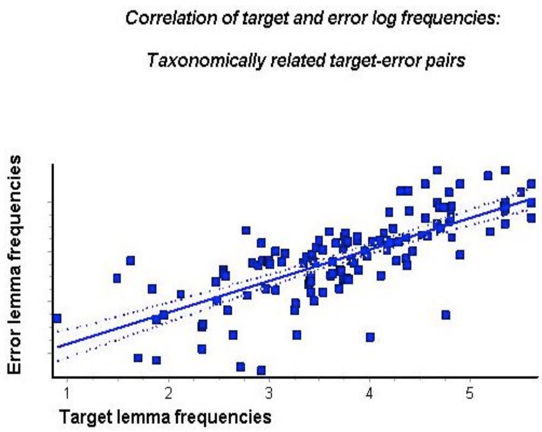
Paired t-tests were run to compare **log-transformed values** of target and error lemma frequencies in each of the two groups of semantic substitutions, taxonomically related and non-taxonomically related pairs. For either group, the frequency difference appeared to be very statistically significant: Within *taxonomically related* group, t (135) = 2.644, p = 0.0092 for lemma frequencies, and t (161) = 2.84, p = 0.0051 for lexeme frequencies; within *non-taxonomically related* group, t (100) = 4.246, p < 0.0001 for lemma frequencies, and t (102) = 3.923, p = 0.0002 for lexeme frequencies. A paired t-test run within *overall* semantic substitutions also reveals an extremely statistically significant target-error frequency difference: t (237) = 4.622, p < 0.0001 for lemma frequencies, and t (264) = 4.639, p < 0.0001 for lexeme frequencies (the number of lexeme pairs analysed is bigger than the number of lemma pairs because there are more duplicate examples of target-error lemma pairs).

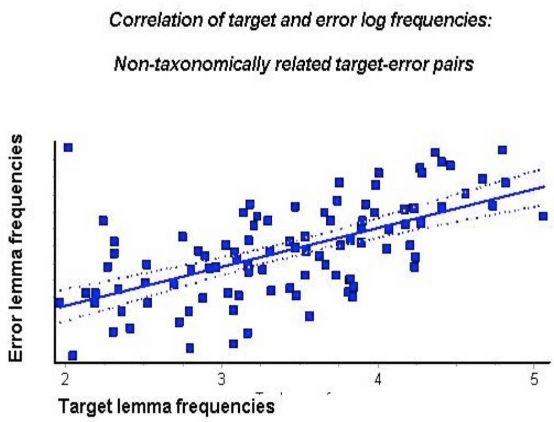




3. Target frequencies are positively correlated with error frequencies.

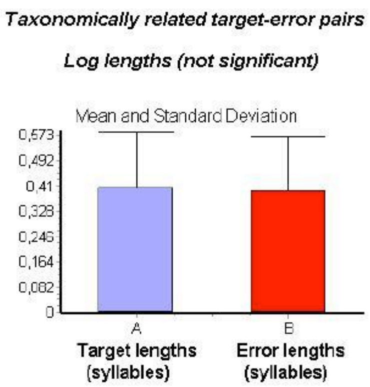
A linear regression test reveals a strong correlation between target and error lemma frequencies (for *overall* semantic substitutions, r = 0.7089, p < 0.0001; for *taxonomically related* target-error pairs, r = 0.7434, p < 0.0001; for *non-taxonomically related* pairs, r = 0.6031, p < 0.0001).

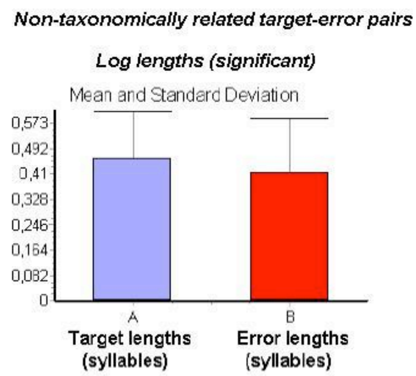




4. Targets that elicit taxonomically related errors do not differ in length from their substitutes; targets that elicit non-taxonomically related errors do (their substitutes tend to be shorter than the targets).

Since the log-transformed length values do not appear to follow a Gaussian distribution, a Wilcoxon matched-pairs test was run to compare the target lengths to the error lengths. Within *overall* semantic substitutions, the difference in length (measured in **syllables**) between targets and their substitutes is not quite significant (W (sum of ranks) = 1537.0, p = 0.0771, mean target log length = 0.4269, mean error log length = 0.4010). Likewise, a Wilcoxon test run to compare the log lengths of targets and their substitutes within the *taxonomically related* group showed no significant difference (W = 131.0, p = 0.7324, mean target log length = 0. 0.4012, mean error log length = 0.3923). However, there appears to be a significant difference in length between targets and errors within the *non-taxonomically related* group (W = 533.0, p = 0.0287, mean target log length = 0.4616, mean error log length = 0.4127).





Similarly, target and error lengths measured in **phonemes** differ significantly for *non-taxonomically related* target-error pairs (W = 1063.0, p = 0.0124, mean target log length = 0.8633, mean error log length = 0.8174), do not differ for *taxonomically related* pairs (W = 769.00, p = 0.3608, mean target log length = 0.7997, mean error log length = 0.7882), and differ for *overall* semantic substitutions (W = 3987.0, p = 0.0228, mean target log length = 0.8244, mean error log length = 0.7995, but the difference is likely to be due to the difference in the non-taxonomically related group).

5. Target length affects error frequency when target and error are non-taxonomically related; however, it is not a significant factor in determining the outcome of the error when target and error are taxonomically related.

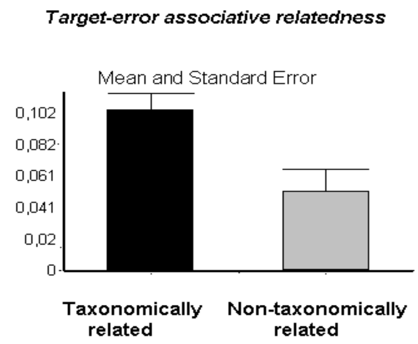
A multiple regression test shows that within *non-taxonomically related* group, both target lemma frequency and target length (measured in **syllables**) are significant variables in predicting error lemma frequency (R squared = 39.82%, $p < 0.0001$; target lemma frequency: $p < 0.0001$, target lexeme length: $p = 0.025$) while for *taxonomically related* pairs, the error frequency is determined by target frequency but not by target length (R squared = 57.06%, $p < 0.0001$; target lemma frequency: $p < 0.0001$, target lexeme length: $p = 0.9170$). Within *overall* semantic substitutions, target lemma frequency appears to be a significant variable whereas target lexeme length is not (R squared = 51.45%, $p < 0.0001$; target lemma frequency: $p < 0.0001$, target lexeme length: $p = 0.1849$).

However, target length measured in **phonemes** is not a significant variable in predicting error frequency either within *overall* semantic substitutions ($p = 0.2647$) or within any of the two groups (for *taxonomically related* pairs, $p = 0.8190$; for *non-taxonomically related* pairs, $p = 0.0968$, not quite significant).

6. The associative links between the target and its substitute are more likely to be stronger when they are taxonomically related than when they are not.

The target-error pairs were analysed in terms of word association norms from the Russian Word Association Thesaurus (which includes 7,000 stimulus words and 105,000 associative response words and phrases). Since the associative responses for the Thesaurus were collected using a free association test, the available responses are extremely variable, so their relative frequencies are often fairly low. The numbers of target-error pairs analysed are smaller here than the total available numbers of examples because not all the target words are included in the Word Association Thesaurus.

For the purpose of the present study, a target word was assumed to be a stimulus word and the substitute word, to be its associative response. Since the raw response frequency values do not follow a Gaussian distribution and square root transformation failed to normalize the distributions (log transformation is not appropriate as there are too many zero response frequency values, especially in the non-taxonomically related group), a nonparametric Mann-Whitney test was used to compare the two groups of semantic substitutions in terms of target-error associative response frequency. The test shows that the frequency is significantly higher for *taxonomically related* target-error pairs compared to *non-taxonomically related* pairs (for square root-transformed frequency values, Mann-Whitney U-statistic = 1873.0, $p = 0.0004$, mean sq. root frequency for *taxonomically related* pairs = 0.1044; mean sq. root frequency for *non-taxonomically related* pairs = 0.0520).



7. Taxonomically related target-error pairs are much more likely to co-occur in speech and texts than non-taxonomically related pairs.

The Russian National Corpus (which contains about 52,000 texts and transcripts totaling 150,000,000 tokens) was used to estimate the mutual informativeness, or co-occurrence strength, of the target and the substitute in the two groups of semantic substitutions.

Mutual information (MI) and T-score were used as the measures of lemma co-occurrence strength. MI compares the frequency of co-occurrence of two words with the frequency of their independent occurrence, whereas T takes into account mainly the absolute frequency of co-occurrence of two words. Since MI is known to overestimate low-frequency words, T-score was used in addition to MI because it highlights the word pairs whose co-occurrence frequency is high enough to be reliable. Thus, the MI test measures the strength of co-occurrence, whereas the T-test measures the confidence with which the co-occurrence frequency can be claimed to be reliable. Two words were regarded as mutually informative (i.e. the occurrence of a word in a sentence increases the probability of occurrence of the other word of the pair) when both MI and T-scores exceeded their minimum acceptable values: MI ≥ 3 and T-score ≥ 2 (see Stubbs 1995; Hunston 2002). MI and T-scores were computed for each target-error pair with a context window of $\pm$ 10 (the average length of a Russian language sentence) by the following formulas (Church, Gale, Hanks, & Hindle 1991):

$$MI(n,c) = \log_2 \{[f(n,c) \times N] / f(n)f(c)\}$$

$$T = \{[f(n,c) / N] - [f(n)f(c) / N^2]\} / \{[\sqrt{f(n,c)}] / N\},$$

where n and c are the two cooccurring words (designated as ‘node’ and ‘collocate’, respectively), N is the size of the corpus (in word tokens), $f(n)$ is the frequency of the first word of the pair (node), $f(c)$ is the frequency of the second word (collocate), and $f(n,c)$ is the cooccurrence frequency of the two words.

The MI and T-scores for sample lists of taxonomically related and non-taxonomically related target-error pairs are shown in Tables 1 and 2 (next page), respectively.

Table 1. MI and T-scores for sample taxonomically related target-error pairs

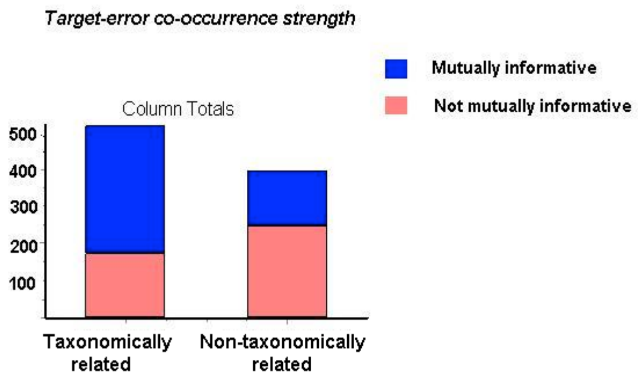
Target	Error	MI	T-score
nedelja *week*	mesjac *month*	3.18	28.67
dnja *day*	noči *night*	2.77	74.86
ogurcov *cucumbers*	pomidorov *tomatoes*	9.79	15.41
čaj *tea*	sok *juice*	2.65	6.18
bljudce *saucer*	tarelku *plate*	6.49	4.84
avtobus *bus*	poezd *train*	3.80	8.76
varežki *mittens*	perčatki *gloves*	8.01	3.59
užin *supper*	zavtrak *breakfast*	6.39	15.15

Table 2. MI and T-scores for sample non-taxonomically related target-error pairs

Target	Error	MI	T-score
vesnoj *in spring*	v subbotu *on Saturday*	-0.62	-1.07
volosy *hair*	golovu *head*	3.44	42.62
želaniem *wish*	udovol′stviem *pleasure*	1.63	7.14
zimoj *in winter*	nočju *at night*	1.49	10.50
magazin *foodstore*	xolodil′nik *fridge*	2.68	3.78
rezinku *eraser*	probku *cork*	0	0
slučaj *case*	variant *version*	0.99	7.55
ekzameny *exams*	kanikuly *vacation*	4.39	3.81

Since both log and square root transformations failed to normalise the MI and T distributions, nonparametric Mann-Whitney tests were run to compare the lemma co-occurrence measures (MI and T-scores) of target-error pairs in the two groups of semantic substitutions. The results of the tests reveal that the values of both co-occurrence measures are much higher for taxonomically related pairs compared to non-taxonomically related pairs (**MI scores**: Mann-Whitney U-statistic = 4408.0, $p < 0.0001$, mean MI for *taxonomically related* pairs = 4.794, mean MI for *non-taxonomically related* pairs = 3.213; **T-scores**: Mann-Whitney U-statistic = 3289.0, $p < 0.0001$, mean T-score for *taxonomically related* pairs = 14.357, mean T-score for *non-taxonomically related* pairs = 4.839).

A Fisher's exact test was run to compare the two groups of semantic substitution errors, taxonomically related vs. non-taxonomically related target-error pairs, in terms of the numbers of mutually informative target-error pairs. An analysis of a 2 x 2 contingency table shows that the rate of mutually informative target-error pairs ($MI \geq 3$ and $T \geq 2$) is considerably bigger in the *taxonomically related* group, and the difference is extremely statistically significant ($p < 0.0001$, odds ratio = 3.351).



Discussion

The results show that higher-frequency nouns tend to substitute for lower-frequency semantically related nouns, with the frequency of a substitute being significantly higher than that of the target. The tendency seems to be more pronounced for lower-frequency nouns that elicit non-taxonomically related errors but it also holds for higher-frequency nouns that elicit taxonomically related errors (although the effect is not observed with raw frequency values, it is very significant with log transformed values). The positive correlation between target frequencies and error frequencies within overall semantic substitutions indicates that target word frequency affects the outcome of the error: to successfully compete with a word to be produced, a semantically related word has to be of a higher frequency. The data suggest that probabilistic information about lexical units may be available at earlier stages of word retrieval such as the stage of lemma retrieval. Thus, the word frequency effect appears to play a role throughout the process of lexical selection. This finding runs counter to the claim that frequency is coded only at the lexeme (phonological encoding) level.

With low-frequency nouns that elicit non-taxonomically related errors, both target frequency and target length (measured in syllables) appear to have an impact on the outcome of the error (the effect of word length measured in phonemes is questionable, suggesting that in Russian, syllable word length might be a more reliable measure). In this case, the error is typically an associatively related word belonging to a cognitive frame evoked by the target, and is often (though not necessarily) at the periphery of the target word's semantic field.

Conversely, higher-frequency nouns that elicit taxonomically related errors are "matched for length" with their substitutes, so word length is not a significant variable predicting the outcome of taxonomically related substitutions. The errors (category coordinates) elicited by such nouns seem to be more closely related to their targets, i.e. have higher measures of associative relatedness and co-occurrence strength, so they are likely to be closer to the centres of the targets' semantic fields. With this type of substitutions, the closeness of the target-error associative links and the frequency of the experience of using both words together in speech and texts might outweigh the significance of the word length factor. The results point to a special psychological status of taxonomic relationships in a word's semantic field, suggesting that, on a paradigmatic scale, the mutual informativeness of lexical items may play a role in forming meaning relationships within a semantic category.

Conclusion

In sum, evidence from semantic paraphasias in normal speech provides support for the hypothesis that words' abstract representations compete for selection and that word frequency is one of the factors which determine the outcome of the competition. Word length is another significant factor when the target competes with a non-taxonomically related item; in semantic substitutions of taxonomic type, the effect of word length seems to be outweighed by the speaker's experience of frequently using the target word together with the competing taxonomically related item (resulting in stronger associative links between the two words).

References

Church, K., Gale, W., Hanks, P., & Hindle, D. 1991. Using statistics in lexical analysis. *Lexical Acquisition*, ed. U. Zernik, 115-64. Englewood Cliff, NJ: Erlbaum.

Damian, M.F. & Als, L.C. 2005. Long-lasting semantic context effects in the spoken production of object names. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 31: 1372–84.

Del Viso, S., Igoa, J. M., & Garcia-Albea, J. E. 1991. On the autonomy of phonological encoding: Evidence from slips of the tongue in Spanish. *Journal of Psycholinguistic Research* 20: 161–185.

Finocchiaro, C., & Caramazza, A. 2006. The production of pronominal clitics: implications for theories of lexical access. *Language and Cognitive Processes* 21: 141–180.

Gerhand, S., & Barry, C. 2000. When does a deep dyslexic make a semantic error? The roles of age- of-acquisition, concreteness, and frequency. *Brain and Language* 74: 26–47.

German, D. J., & Newman, R. S. 2004. The impact of lexical factors on children's word-finding errors. *Journal of Speech, Language, and Hearing Research* 47: 624–636.

Gordon, J. K. 2002. Phonological neighborhood effects in aphasic speech errors: Spontaneous and structured contexts. *Brain and Language* 82: 113–145.

Harley, T. A., & MacAndrew, S. B. G. 2001. Constraints upon word substitution speech errors. *Journal of Psycholinguistic Research* 30: 395–418.

Hotopf, W. H. N. 1980. Semantic similarity as a factor in whole-word slips of the tongue. *Errors in linguistic performance: Slips of the tongue, ear, pen and hand*, ed. V. Fromkin, 97–109. New York: Academic Press.

Hunston, S. 2002. *Corpora in Applied Linguistics*. Cambridge: Cambridge University Press.

Jescheniak, J.D., & Levelt, W.J.M. 1994. Word frequency effects in speech production: Retrieval of syntactic information and of phonological form. *Journal of Experimental psychology: Learning, Memory, and Cognition* 20(4): 824-843.

Jurafsky, D. 2003. Probabilistic modeling in psycholinguistics: Linguistic comprehension and production. *Probabilistic Linguistics*, eds. R. Bod, J. Hay, and S. Jannedy, 39–95. Cambridge, MA: MIT Press.

Karatsinski, V. 2010. Frequency of use leads to automaticity of production: Evidence from repair in conversation. *Language and Speech* 53(1): 71–105.

Kittredge, A. K., Dell, G. S., Verkuilen, J., & Schwartz, M. F. 2008. Where is the effect of frequency in word production? Insights from aphasic picture-naming errors. *Cognitive Neuropsychology* 25: 463–492.

Kroll, J.F., & Stewart, E. 1994. Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language* 33: 149–74.

Navarette, E., Basagni, B., Alario, F.-X., & Costa, A. 2006. Does word frequency affect lexical selection in speech production? *The Quarterly Journal of Experimental Psychology* 59: 1681–1690.

Oldfield, R. C., & Wingfield, A. 1965. Response latencies in naming objects. *The Quarterly Journal of Experimental Psychology* 17: 273–281.

Russian Word Association Thesaurus/Русский Ассоциативный Словарь. 2002. Ред. Ю. Н. Караулов, Г. А. Черкасова, Н. В. Уфимцева, Ю. А. Сорокин, Е.Ф.Тарасов. Москва: Астрель.

Stubbs, M. 1995. Collocations and semantic profiles: on the cause of the trouble with quantitative studies. *Functions of Language* 2, 1: 23–55.

Vitevitch, M. S. 1997. The neighborhood characteristics of malapropisms. *Language and Speech* 40: 211–228.