University of Pennsylvania Working Papers in Linguistics

Volume 7
Issue 2 *Current Work in Linguistics (Working Papers)*

Article 6

1-1-2002

Centering in Russian

Sophie A. Malamud
University of Pennsylvania, smalamud@sas.upenn.edu

Centering in Russian

Centering in Russian*

Sophia A. Malamud

1 Introduction

The management of a hearer's (reader's) attention is an integral part of cooperative communication in any language. Discourse is thus structured in a way that allows the hearer to focus his attention on various entities evoked, and to ensure that information about them is entered into his knowledge-store in a coherent way.

Following Prince (1988) and Vallduví (1990), among others, I shall be referring to this non-truth-conditional component of sentence-meaning as 'informational component' of a sentence. Packaging this meaning into syntactic structures will be termed the 'information packaging' or 'informational structure' of a sentence (*ibid*). A part of this activity is concerned with focussing the hearer's attention on a single entity in each sentence (the topic in the sense of Givón 1983 and Miltsakaki 1999). On the extra-sentential level, navigating the hearer's attention from one topic to the next determines attentional structure of discourse.

The goal of this paper is to investigate the attentional and informational structure in Russian written narrative, in order to shed some light on the question: What are the principles determining the change and maintenance of local topics in Russian discourse?

This paper will proceed as follows: in Section 2, I will give a concise overview of the theoretical tools used in the present study. Section 3 will present a corpus study of Russian written narrative which attempts to answer the question of the investigation. Section 4 will provide some discussion of the results, with a brief conclusion following in Section 5. The Appendix will then provide a worked example and references.

*My deepest gratitude is due to the people who have shaped my undergraduate encounter with linguistics: professors and teaching assistants in Linguistics and Cognitive Science. I want to thank especially Ellen Prince, who has inspired and guided my growth as a researcher. I am also grateful to Eleni Miltsakaki for lengthy and productive discussions of data and theory. Needless to say, all shortcomings and mistakes remain my own.

2 Centering Theory: Information Structure and Entity-tracking

For the purposes of this paper, I shall utilize the Centering Theory (Grosz, Joshi and Weinstein 1995, Brennan, Friedman, and Pollard 1987, Walker, Iida, and Cote 1994) to provide an algorithmic definition of 'topic' (Mathesius 1915, Hockett 1958, Strawson 1964, Gundel 1974, 1988, Kuno 1980, *inter alia*), more or less equivalent to Halliday's 1967 'theme' or Vallduví's 1990 'link'. The topic/theme is used as an address for the hearer under which to enter new information (comment) in his mental knowledge store. I will not require this topic to be sentence-initial, contra Halliday (1967) ('theme') and Vallduví (1990) ('link'). Dropping this requirement allows us to explore the relationship between topic-structure and word order.

Joshi and Kuhn (1979) and Grosz, Joshi and Weinstein (1995) develop an entity-based model of local attention structure. By centering a single semantic entity in an utterance and predicating over it, the theory constrains the inferencing required for such tasks as anaphora resolution; computed transitions from one center to the next model the local discourse coherence.

Centering Theory has been proposed as such in Grosz, Joshi, and Weinstein (1983) (published as Grosz, Joshi and Weinstein 1995); the ideas were subsequently developed and expanded both by the original authors, and by others (Brennan, Friedman, and Pollard 1987; Walker, Iida, and Cote 1994; Walker, Joshi, and Prince 1998, *inter alia*). The theory provides the most operational definition of a topic so far. It will therefore be utilized here as a formal framework for exploring the topic-structure of written Russian discourse. The basic notions of the theory are defined and discussed below.

2.1 The Centering Transitions

The Centering Algorithm allows us to compute the smoothness of transition between utterances based on a salience ranking of entities in a discourse. In this section, I discuss the basic definitions needed to understand Centering Theory, as well as the different transition types.

Definition 1: For each utterance, the set of discourse entities evoked in it constitutes the set of *forward-looking centers* (Cf). Centers are semantic entities that are part of the discourse model (see Heim 1983), or items in the set of shared current concern for speaker and hearer (Yokoyama 1986).

Definition 2: There is a special member of this set called the *backward-looking center* (Cb). This is the entity that is most central in the utterance (Walker and Prince 1996), the file card you're writing on (Reinhart 1982,

Heim 1983), corresponding to 'the utterance topic/theme' (Kuno 1980, Reinhart 1982). The Cb is the entity which links the current utterance with the previous discourse.

The set of forward-looking centers is ranked according to discourse salience, or 'activatedness'. The factors that determine ranking are the crux of the Centering Algorithm. By definition, if any centers of the current utterance are evoked in the subsequent utterance, the highest-ranked one is the Cb of that subsequent utterance. In fact, the following Pronoun Rule has been formulated, and reflects the observation that the most cognitively prominent entities should need the least description for successful reference: 'If there is a pronoun in an utterance, then the Cb of this utterance is also denoted by a pronoun' (Grosz, Joshi and Weinstein 1995).

Definition 3: The highest-ranked center is the *preferred center* (Cp). It predicts what the next utterance is going to be about.

The interaction between Cb and Cp determines smoothness of transition from one utterance to the next as shown in Table 1 below. When the most central entity in an utterance ($Cb(U_n)$) is the same as the most central entity in the previous utterance ($Cb(U_{n-1})$), and the same item is also predicted to be central in the next utterance ($Cp(U_n)$), the resulting discourse is very coherent, and the transition is *Continue*. On the other hand, when the Cb from a previous utterance is retained as such, but not predicted to be as salient in the next utterance, the transition type is *Retain*. The two Shifts result when the most central entity changes: the *Smooth-Shift* predicts that it should not change again in the next utterance, while the *Rough-Shift* does (Table 1).

Table 1: Transitions from U_{n-1} to U_n[1]

	$Cb(U_n) = Cb(U_{n-1})$	$Cb(U_n) \neq Cb(U_{n-1})$
$Cb(U_n) = Cp(U_n)$	Continue	Smooth-Shift
$Cb(U_n) \| not = Cp(U_n)$	Retain	Rough-Shift

The transitions, smoothest to roughest, are: Continue, Retain, Smooth-shift, and Rough-shift (Walker and Prince 1996). Centering analyses have shown that smoother transitions are preferred over rougher ones within a discourse segment (Di Eugenio 1998, Rambow 1993, *inter alia*).

[1] U_n is the n^{th} utterance; $Cb(U_n)$ is the backward-looking center of the n^{th} utterance; and $Cp(U_n)$ is the preferred center of the n^{th} utterance.

2.2 The Ranking

The ranking of entities determines the Cp of the current utterance, and predicts the Cb of the next one. The ranking principle arrived at by most Centering analyses (e.g. Di Eugenio 1998, Miltsakaki 1999) is based on the grammatical function of the entities, which are ranked as follows: EMPATHY → SUBJECT → OBJECT → OTHER. Here, 'empathy' denotes phrases grammatically marked as 'empathic' (e.g. in Japanese); the grammatical category was later transformed into a semantic one to include constructions in other languages clearly emphasising the experiencer (e.g., in the dative subject constructions, see Yokoyama 1986).

Studies of Italian (Di Eugenio 1998), Turkish (Hoffman 1998), and Greek (Miltsakaki 1999) have shown that this ranking indeed correctly predicts full noun phrase, pronoun, and zero-pronoun usage, and is independent of the utterance word order in these languages.

However, a study of German (Rambow 1993) showed that whereas topicalization interacts with Centering in an ambivalent way, scrambling in the Mittelfeld directly affects the ranking: "the Cf (ordered set of forward-looking centers) of an utterance is the list of constituents of the Mittelfeld in that order." Thus, in German, re-arranging constituents that follow the inflected verb (V2) will change the center (topic) of the next utterance, and affect the local coherence of discourse.

2.3 The Segment, the Utterance, and Other Ranking Assumptions

Centering is a model of local discourse structure and operates within discourse segments. Hence, it is important to know how to determine the segmentation of a discourse. However, determination of segment boundaries is a separate question of much current investigation. For the purposes of this study therefore, I assume no *a priori* segmentation in written discourse outside of the writers' segmentation of large books into chapters.

Within each segment, the Centering algorithm calculates the Cf list for every 'utterance', which is another notion in need of formal definition. Early Centering analyses seem to assume the utterance to be approximately the tensed clause (Kameyama 1998). In a later investigation (Miltsakaki 1999), this was revised, and 'utterance' was defined as a full sentence, i.e. "the main clause and its accompanying subordinate and adjunct clauses" (Miltsakaki and Kukich 2000). I follow here this revised definition. Miltsakaki (1999) argues that the ordering of subordinate and main clauses does not affect Centering. In this study, therefore, unless there were two or more coordinated subordinate

clauses, nothing outside the main clause had significant effect on the ranking of the Cf-list.

There has been much variation as to the correct ranking of entities within a complex noun phrase (e.g. possessives). In my corpus, the proportion of possessive noun phrases for which the various theories would predict different rankings is negligible, and so one may safely assume here the simple principle of left-to-right ranking.

3 Centering Study of Russian

3.1 Pronoun Rule

Recast in the terms of Centering theory, the question "What are the principles determining the change and maintainance of local topics in Russian discourse?" may be reformulated as follows: What is the ranking principle for the list of forward-looking centers in Russian? The two main possibilities for the ranking are the ones already stated for other languages: ranking by grammatical function or ranking left-to-right in the main clause. In Russian, a language with flexible word order which is canonically SVO, these two principles would rank the entities differently only in sentences with OVS, OSV, or VOS word order.

A corpus study of written Russian was conducted to investigate which ranking principle holds in Russian. The main source of data in this study was the online library of Russian literature (*www.lib.ru*). I have chosen a number of literary narrative segments containing scrambled sentences. A computerized search was used to select the segments containing scrambled sentences from electronic books. A total of 44 analysable segments of two or more sentences were found, each containing at least one scrambled sentence.

In an attempt to determine the ranking principle for Russian, the Centering Theory's Pronoun rule was utilized. The corpus was searched for sentences in which of the two or more possible Cb candidates only one was pronominalized. There were 16 such sentences in the data (cf. 127 total Cbs). For these backward-looking centers ($Cb(U_n)$), the word order of the preceding utterance (U_{n-1}) and the grammatical role of the expression in U_{n-1} referring to the $Cb(U_n)$ were traced. The results are summarized in Table 2 below.

As is evident from Table 2, the 4th, 5th, 8th, 9th, and 10th (maybe also 7th, 11th and 13th) tokens speak in favour of the grammatical-function-based ranking, while the 14th and possibly 6th and 7th work left-to-right. However, the number of tokens is too small for any significant or definite conclusion,

especially because, first, there are no examples with OVS or VOS word order in U_{n-1}, second, ranking preference for some tokens depends on the assumption that a semantic EMPATHY overranks a subject, and third, the tokens 8, 9, and 10 form a parallel-construction sequence in a paragraph. Thus, I turn to a statistical study of the corpus, hoping that tracing the actual entities chosen for centering by the narrators will provide a clearer picture of their salience in preceding discourse.

Table 2. The Pronoun Rule study[2]

#	Word order in U_{n-1}	Grammatical role of $Cp(U_{n-1})$	Ranking principle favoured
1	SVO	subject	both
2	SVO	subject	both
3	SVO	subject	both
4	OSV	subject	G
5	OSV	subject	G
6	DativeVCompl(SV)	EMPATHY (dative experiencer)	both or W
7	S-prep-GenitiveV	EMPATHY (genitive experiencer)	G or W
8	OSViO	subject	G
9	OSViO	subject	G
10	OSViO	subject	G
11	InstrumentalVS	subject	G
12	SVCompl(VS)	subject of the complement clause	
13	pPrepositionalVS	subject	G
14	iOSVO	indirect object	W
15	SVO	subject	both
16	SVO	subject	both

3.2 Comparison with Non-scrambling Text

3.2.1 Control Data

To provide a measure of the true proportions of different transitions in Russian texts, a full short story "Pyat' minut vzajmy" was chosen and analysed. A

[2]G: ranking by grammatical function; W: ranking by word order.

total of about 70 transitions were calculated. Since of the 78 sentences containing 24 transitive clauses with overt arguments only four were scrambled, the ranking was performed by grammatical function only. Discounting the rough-shifts in the opening and closing paragraphs of the story as 'necessities of artistic considerations,' the analysis reveals that Rough-Shifts constituted 10% of all the transitions, with the remaining comprising 34 Continues (48.5%), 16 Retains (23%), and 13 Smooth-Shifts (18.5%).

3.2.2 The Rough-Shift Measure

Twenty native Russian speakers, representing a wide variety of ages (17-74 years), occupations (students, sociologists, physicists, computer technicians, businessmen, beauty salon workers, translators, housewives, and a mover), and geographical backgrounds (Moscow, Kharkov, Kiev, Leningrad, various suburbs in Russia and the Ukraine, and several immigrants in Brooklyn, New York) were chosen as informants. They were asked to read both the control story and a portion of the scrambled data corpus matched for size (ten read the control first, and ten the scrambled portion first). At the end, the informants were asked to rate the 'perceived coherence' of randomly chosen discourse segments. The reading times and the ratings were recorded. Although some individual variation in reading speed was detected, there was no significant difference in reading times between the control and scrambled data. The ratings indicated that the scrambled corpus was perceived as slightly less coherent, which was attributed to its fragmentary nature: "You have to switch your brain from one excerpt to the next," as one of the informants put it.

The Centering analysis of this data was done manually twice (see Appendix A). The first analysis utilized ranking by grammatical function and produced 50 Continues, 46 Retains, 16 Smooth-Shifts, and 17 Rough-Shifts out of 129 total transitions. Then, using the left-to-right ranking hypothesis, the second analysis was performed, producing 49 Continues, 46 Retains, 21 Smooth-Shifts, and 13 Rough-Shifts.

In their 2000 Centering study, Miltsakaki and Kukich argue that "in general, Continues, Retains, and Smooth-Shifts do not yield incoherent discourses" (Miltsakaki and Kukich 2000). Therefore, only the presence of a Rough-Shift signals a significant incoherence. The presence of Rough-Shifts in a perceptually coherent discourse and the number of Rough-Shifts were therefore the first considered factors in this study.

Statistical tests were run on these numbers, with the transition percentages from the short story analysis serving as controls, i.e., the norm. Although the tests indicated that the second analysis was much closer to the normal

data, the sample was not large enough to yield a degree of certainty above 75%. Meanwhile, a qualitative evaluation of data was performed. A closer examination of the transitions indicated that of the 17 Rough-Shifts produced by the first analysis (by grammatical role), 6 were found to be Smooth-Shifts in the second (by linear order). One of these could have been a Continue changing to Retain in the second analysis, depending on the judgement of the main clause boundaries. The remaining 11 were Rough-Shifts in the second analysis as well.

At the same time, out of the 13 Rough-Shifts produced by the second analysis (by linear order), 2 were Smooth-Shifts in the first analysis (by grammatical role). One of these could have been actually a Smooth-Shift in the second analysis if a different ranking principle for the complex noun phrases were adopted.

Based on the Rough-Shift measure alone, we can conclude that word-order dependent ranking provides a more accurate measure of discourse coherence than the grammatical-function based one. However, both rankings perform poorly on scrambled data: while the informants' judgements and reading times indicated that this text was similar to control data, the distribution of transition percentages was very significantly different for the two. The chi-squared test resulted in less than 1% probability that the difference is due to chance, for both ranking hypotheses (see Table 3 in the next section).

3.2.3 New Hypothesis: Incorporating the Verb

The two rankings discussed above (by grammatical role or by word order) have produced approximately the same number of Continue and Retain transitions. Moreover, both analyses have 'improved' and 'worsened' about the same number of these transitions. This suggests that neither hypothesis sufficiently accounts for the more coherent data.

In the original left-to-right ranking hypothesis, no consideration has been given to the verb. However, it has been noted for many languages, including Russian, that the pre-verbal and post-verbal positions in an utterance have different informational functions (Yokoyama 1986, Rambow 1993, Kiss 2000, *inter alia*). Therefore, the position of the verb was traced in the 34 scrambled transitive sentences with overt arguments for which the two analyses give different transitions. For 15 of them the word-order dependent ranking (second analysis) produced a smoother transition, whereas for the remaining 19, the other analysis did.

Crucially, 12 of the former sentences had OVS and the remaining 3 had VOS word order, whereas 16 of the latter had the order OSV. The remaining 3

sentences contradicting the left-to-right hypothesis were OVS. However, two of them were a part of the 6-utterance parallel construction segment, and one more a part of a segment in which calculation of segment boundaries and, therefore, of the Cbs, was very difficult. Thus, it becomes obvious that simply scrambling the object to the sentence-initial position in Russian doesn't affect its discourse salience, but serves some other purpose. When, however, the subject is demoted to the post-verbal position, the salience of both entities is affected. The corpus contained no instances of the VSO word order; thus it is unclear whether post-verbal status of the subject is sufficient to demote it in the ranking, or whether it must follow the object as well.

The new ranking hypothesis is formulated as follows: The entities in Cf are ranked by grammatical function, unless the subject is in the post-verbal position. This revised hypothesis was used for the third and final analysis of the data. The analysis produced 11 Rough-Shifts, 20 Smooth-Shifts, 35 Retains, and 63 Continues. These are the smoothest resulting transitions yet. The chi-squared test was used to measure the probability that observed differences in all the variables (the number of occurrences of each transition type) are the result of chance variation. Low chi-square values indicate that the distribution of transitions is essentially the same as in the control data set. In this test, the significance of the Rough-Shift measure is somewhat downplayed, since each variable is given the same significance in the calculation of the chi-squared value. Again, the percentages from the short story analysis were used as controls. As is evident from Table 3, the new hypothesis results in a significantly more normal analysis of the scrambling data (60% probability that the difference from normal is chance).

Table 3. The chi-squared test

Ranking hypothesis	Chi-squared value	Probability
By word order	12.69	Less than 1%
By gram. function	16	Even less! (about 0%)
The new hypothesis	2.07	60%

Thus indeed, the scrambling that places the subject in clause-final position reduces the salience (topicality) of the subject and predicts an earlier element to be centered in the subsequent utterance. Possible reasons for this peculiar ranking principle are considered below in Section 4.

4 Discussion: Intonation and Information

The ranking principle arrived at in the previous section indicates that occurring sentence-finally corresponds to the lowered salience of an item, prohibiting it from being the preferred center. This coincides with the observation made by Yokoyama (1986) that the items outside the hearer's knowledge set are placed last in the sentence, immediately following the items outside the set of hearer's current concern. Yokoyama has formulated corresponding ordering rules for Russian, partially formalising the Prague school's 'theme-rheme condition' (roughly the given-to-new ordering of constituents, see interpretation in Rambow 1993).

Yokoyama's rule was formulated for the Type I intonation contour. According to Svetozarova (1998), this intonation contour "can be called *neutral* ... Neutral sentence stress at the end of a final declarative sentence is characterized by a simple falling tone and increased length of the stressed vowel with relatively low intensity." (Svetozarova 1998:266). A number of researchers have argued that the element bearing the sentential stress in the Type I contour is thereby marked: "A falling nuclear accent (HL*) corresponds to the natural focus. The exponent of natural focus in Russian is constituted by the last lexical accent, i.e. at the right periphery of a sentence" (Zybatow and Mehlhorn 1999, cf. Bryzgunova 1971, Krylova and Khavronina 1988).

Following studies on English (Steedman 2000), and Italian (Cinque 1993), and noting the claims by Zybartow and Mehlhorn that natural focus is at the right periphery, I hypothesize that the lower salience for the sentence-final item is dependent not on its position as such, but rather on the sentence-final neutral sentential stress in the sentences in question. This provides an explanation for the effect of word order on salience ranking of entities in discourse.

5 Conclusions

The Centering study of a written corpus suggests that word order and attentional structure of discourse are interdependent phenomena in Russian. The entity-based approach to local discourse coherence shows a special informational status of subject-final word orders in Russian. The study shows that in such word orders the subjects are dispreferred as potential topics for subsequent discourse. As previous studies suggest, this may be due to the informational marking induced by the sentence-final sentential stress in neutral Russian intonation.

This study constitutes a step towards understanding the principles of at-

tention management in Russian. A further investigation of spoken discourse is currently under way, as is also an exploration of the reverse influence of information structure on sentential word order in Russian (Malamud 2001).

Appendix: A Comparison of Left-to-Right and Grammatical-Function Rankings

Consider the following segment from Bulgakov, in which the second sentence is scrambled (The Cb of the previous utterance is K.):

(1) K. svistnul.
K. let-out-a-whistle.
'K. let out a whistle.'

Cf = {K.}, Cp = K., Cb = K.

(2) Etogo svista Margarita ne uslyhala, no ona ego uvidela v to
Of-this whistle Margarita not heard, but she it saw at that
vremya, kak ee vmeste s goryachim konem brosilo sazhenej
time, as her together with hot horse it-threw sazhens
na desyat' v storonu.
for ten to side.
'Margarita didn't hear this whistle, but she saw it at the same time when she, together with her hot-tempered horse, was thrown several meters to the side.'

Analysis 1, Ranking by Grammatical Function:

Cf={Margarita, whistle, horse}, Cp=Margarita, Cb = whistle, Transition = Rough-Shift

Analysis 2, Ranking by Word Order:

Cf = {whistle, Margarita, horse}, Cp = whistle, Cb = whistle, Transition = Smooth-Shift

References

Brennan, Susan, Friedman, Marylin, and Pollard, Carl. 1987. A Centering Approach to Pronouns. In *Proceedings of the 25th Annual Meeting of the ACL*, 155-162. Stanford, CA.

Bryzgunova, Elena. 1971. O smyslorazlichitel'nyx vozzmozhnostjax russkoj intonacii [On meaning-distinguishing possibilities of Russian intonation]. *Voprocy jazykoznanija* [Issues in linguistics] 4, 42-52.

Cinque, Guglielmo. 1993. A Null Theory of Phrase and Compound Stress. Linguistic Inquiry 24, 239-297.

DiEugenio, Barbara. 1998. Centering in Italian. In *Centering in discourse*, ed. Walker, Marilyn, Aravind Joshi, and Ellen Prince, 115-138. Oxford: Oxford University Press.

Givón, Talmy. 1983. Introduction. In *Topic continuity in discourse: A quantitative cross-language study*, ed., Givón, Talmy, 5-41 Amsterdam: John Benjamins.

Grosz, Barbara, Aravind Joshi, and Scott Weinstein. 1995. Centering: a framework for modelling local coherence of discourse. *http://www.cis.upenn.edu/ ircs*.

Gundel, Jeanette. 1974. *The role of topic and comment in linguistic theory*. University of Texas PhD dissertation.

Gundel, Jeanette. 1988. Universals of topic-comment structure. In *Studies in syntactic typology*, ed. Hammond, Michael, Edith Moravcsik, and Jessica Wirth, 209-239. Amsterdam: John Benjamins.

Halliday, Michael. 1967. Notes on transitivity and theme in English. Part II. *Journal of Linguistics* 3, 199-244.

Halliday, Michael. 1985. *Introduction to Functional Grammar*. London: Edward Arnold.

Heim, Irene. 1983. File change semantics and the familiarity theory of definiteness. In *Meaning, use and the interpretation of language*, ed. Bauerle, Rainer, Christoph Schwarze, and Arnim von Stechow, 164-189. Berlin: Walter de Gruyter.

Hockett, Charles. 1958. *A Course in Modern Linguistics*. New York: Macmillan.

Hoffman, Beryl. 1998. Word Order, Information Structure, and Centering in Turkish. In *Centering in discourse*, ed. Walker, Marilyn, Aravind Joshi, and Ellen Prince, 250-272. Oxford: Oxford University Press.

Joshi, Aravind and Steve Kuhn. 1979. Centered Logic: The Role of Entity Centered Sentence Representation in Natural Language Inferencing. *6th International Joint Conference on Artificial Intelligence*, 435-439.

Kameyama, Megumi. 1998. Intrasentential Centering: A case study. In *Centering in discourse*, ed. Walker, Marilyn, Aravind Joshi, and Ellen Prince, 89-112. Oxford: Oxford Univ Press.

Kiss, Katalin. 2000. Movement to the left periphery. Ms. Linguistic Institute of the Hungarian Academy, *International Conference on Word Order and Scrambling*, Tuscon, AZ.

Krylova, Olga and Serafima Khavronina. 1988. *Word order in Russian.* Moscow: Russkij Yazyk.

Kuno, Susumu. 1980. The scope of the question and negation in some verb-final languages. In *Proceedings from the Sixteenth Regional Meeting of Chicago Linguistic Society*, 155-169. University of Chicago.

Malamud, Sophia. 2000. Centering and Scrambling: Toward Discourse-Status Motivation of Russian Word Order. Presented at *The Ninth Annual Workshop on Formal Approaches to Slavic Languages,* University of Indiana, Bloomington.

Malamud, Sophia. 2001. *Attentional and informational structure of Russian: Interaction of reference, coherence, and word order.* University of Pennsylvania undergraduate honours thesis.

Mathesius, Vilem. 1915. O passivu v moderníangličtině. Sborniik filologický 5, 198-220.

Miltsakaki, Eleni. 1999. Dissociating Discourse Salience from Information Structure: Evidence from a Centering study in Modern Greek and Japanese. Presented at *Computational Linguistics in the Netherlands, (CLIN '99).* OTS, Utrecht.

Miltsakaki, Eleni. 2001. *Attention Structure in Discourse: A Cross-Linguistic Investigation.* University of Pennsylvania dissertation proposal.

Miltsakaki, Eleni and Karen Kukich. 2000. Automated Evaluation of Coherence in Student Essays. To appear in *Proceedings of ACL 2000. http://www.ling.upenn.edu/ elenimi/grad.html.*

Maxim Moshkow. On-line Russian Library. *http://www.lib.ru.*

Prince, Ellen. 1988. Discourse analysis: A part of the study of linguistic competence. In *Linguistics: The Cambridge Survey,* vol.2, ed. Newmeyer, Frederick, 164-182. Cambridge: Cambridge University Press.

Rambow, Owen. 1993. Pragmatic Aspects of Scrambling and Topicalization in German: A Centering Approach. Ms., University of Pennsylvania.

Reinhart, Tanya. 1982. Pragmatics and linguistics: An analysis of sentence topics. *Philosophica* 27, 53-94.

Steedman, Mark. 2000. *The Syntactic Process*. Cambridge, MA: MIT Press.

Strawson, Peter. 1964. Identifying reference and truth-values. In *Semantics*, ed. Steinberg, Danny and Leon Jacobovits, 86-99. Cambridge: Cambridge University Press.

Svetozarova, Natalia. 1998. Intonation in Russian. *Intonation Systems. A Survey of Twenty Languages*, ed. Hirst, Daniel and Albert Di Cristo. Cambridge: Cambridge University Press.

Vallduví, Enric. 1990. *The Information Component*. University of Pennsylvania PhD dissertaion.

Walker, Marilyn and Ellen Prince. 1996. A Bilateral Approach to Givenness: a Hearer-Status Algorithm And a Centering Algorithm. In *Reference and referent accessibility*. ed. Fretheim, Thorstein and Jeanette Gundel, 291-306. Philadelphia: John Benjamins.

Walker, Marilyn, Masayo Iida, and Sharon Cote. 1994. Japanese Discourse And the Process of Centering. *Computational Linguistics 21*, 193-232.

Walker, Marilyn, Aravind Joshi, and Ellen Prince (eds). 1998. *Centering Theory in Discourse*. Oxford: Oxford University Press.

Webber, Bonnie. 1991. Structure And Ostension in the Interpretation of Discourse Deixis. *http://www.ling.upenn.edu/ ellen/bonnie.ps*.

Yokoyama, Olga Tsuneko. 1986. *Discourse and Word Order*. Philadelphia: John Benjamins.

Zybatow, Gerhild and Grit Mehlhorn. 1999. Experimental Evidence for Focus Structure in Russian. In *The Eighth Annual Workshop on Formal Approaches to Slavic Languages. The Philadelphia Meeting*, ed. King, Tracy Holloway and Irina Sekerina, 414-434. Ann Arbor: Michigan Slavic Publications.

Department of Linguistics
University of Pennsylvania
Philadelphia, PA 19104-6305
smalamud@sas.upenn.edu