

Structural and Prosodic Correlates of Prominence in Free Word Order Language Discourse

Tatiana Luchkina¹, Jennifer S. Cole¹

¹Department of Linguistics, University of Illinois at Urbana-Champaign, USA

luchkin1@illinois.edu, jscole@ling.illinois.edu

Abstract

Production and perception experiments with native speakers of Russian, a free word order language, show that prosody and change in word order are used to mark discourse-prominent constituents. Concurrent application of these cues to prominence is possible, as evident from distinctively higher f_0 and intensity maxima, and duration values associated with ex-situ words, as well as their higher visibility in discourse. Distinctive acoustic-prosodic realization of ex-situ words may cue their relatively high informational load and discourse prominence, as well as (redundantly) signal that the word is left- or right- dislocated.

Index Terms: Information structure, prosodic prominence, word order

1. Introduction

Independently of the modality of presentation, human processing of discourse involves identifying the information structure status of discourse elements. Information structure, and the related notion of accessibility, have been offered a variety of interpretations in the linguistic and psychological literature [1,2,3]. One approach adopts a tripartite distinction of discourse entities into the most accessible (or given, active), least accessible (or novel, inactive) and inferable (semi-active) [4]. Discourse-novel entities are described as more prominent than the discourse-given ones. New information (also known as focus or rheme), is opposed to discourse-given information, (topic or theme) as categories of information structure (IS) and are differentiated with regards to their visibility in discourse or *discourse prominence* [5].

Prosody, morphology, and the structural organization of information provide different means of encoding the discourse status of a word [6,7]. The prosodic encoding of discourse-prominence is a well-studied, psychologically real property of discourse production and perception [8,9,10]. It involves perceptually salient changes in voice quality, duration and intensity, as well as changes in f_0 . The degree of discourse prominence or relative discourse accessibility expressed by prosodic means can be further affected by structural means, i.e., strategic positioning of the discourse-prominent word in a syntactic phrase or clause. Structural prominence is especially suited for the so-called free word order languages, where the syntactic function of a word is marked overtly by means of morphological case. In such languages, the ordering of sentential constituents can be used to encode their relative discourse prominence, and signal their information structure category.

2. Two routes to prominence

Cross-linguistic studies show that languages use *either* prosodic marking or constituent ordering to encode discourse prominence and IS. Based on a comparative study of Italian

and Turkish (relatively free word order languages), and English (fixed word order language), Donati & Nespors [11] propose that languages with rigid word order allow prosodic marking of discourse information ('focus') at different locations in the sentence, while languages with flexible word order do it through word order, and, consequently, exhibit less variation in the location of prosodic prominence. This model, with dual routes for the encoding of discourse information, predicts that it will be relatively uncommon that a language uses both word order and prosodic marking simultaneously to signal novel or important information in discourse. With respect to this [11] suggested to categorize languages into prominence dislocating (e.g., English), i.e., those which utilize prosodic cues to mark prominence, and constituent dislocating (e.g., Turkish or Italian), i.e., those which primarily use word order, or structural cues.

Consider how discourse prominence is expressed in English. While rightmost accent placement is the default location of the prominent constituent in English (1), the phenomenon of metrical reversal or stress shift [12, 13] can displace phrasal prominence leftward to signal the IS marked category of contrastive focus (2):

- (1) Joel bought a green PORSCHE.
- (2) Joel bought a GREEN porsche.

Italian illustrates a different prominence marking strategy [11], where speakers may choose to move the discourse-prominent word to a syntactic location that is systematically associated with its IS category or discourse-prominent status. Thus, while SVO is the canonical word order in Italian (*Mario arrive* 'Mario arrives'), the reverse ordering of S and V confers prominence to the subject (*Arrivo Mario*).

This sort of overt movement is often said to be prosodically motivated, which means that a word undergoes overt movement in order to be associated with the main phrasal prominence position, where it is perceptually salient.

While the proposal by Donati & Nespors parsimoniously accounts for the cross-linguistic distribution of the two distinct prominence encoding mechanisms, structural and prosodic, this work draws attention to the growing body of empirical evidence that word order is an optional resource for encoding IS categories and discourse-prominence in free word order languages, along with prosodic means [14,15,16,17, among others]. Languages which are known to display IS-triggered movement (Spanish, Greek, Russian, Georgian, and Italian, among others) are also known to use prosodic means to mark prominence in-situ, i.e., on constituents that are discourse-prominent, but which have not been moved to a syntactically designated prominent position. For such languages, one interesting question is whether the concurrent application of structural and/or prosodic means is purposeful, i.e., used to encode different categories of information in discourse.

Luchkina and Cole [18] report a preliminary investigation of this issue for Russian, a highly free word order language which simultaneously exhibits structural and prosodic prominence as means of marking information that is novel or

particularly salient. The semantically neutral, default word order in Russian is SVO, and as in other free word order languages, in Russian, a word can appear in its canonical position (in-situ), fronted, or post-posed. The ordering of the constituents in a sentence marks IS and not grammatical function. In the following example from [18], continuations (a) and (b) are both possible for the sentence in (4), but are felicitous under different discourse conditions: given the context provided in (4), the word Ivan, critical to the understanding of who does the cooking, may be located in the rightmost position, where it is structurally prominent (as in b), or may occur pre-verbally as in (a) and receive (optional) prosodic prominence.

(4) Tri druga, Ivan, Petr, i Andrey, nahsli novjij retsept pizzj.
Three friends, Ivan Petr and Andrey, found a new pizza recipe.

a. (Smotri!) IVAN gotovit pizzu.
(look) Ivan-SUBJ cooks pizza-OBJ

b. (Smotri!) Pizzu gotovit Ivan.
(look) pizza-OBJ cooks Ivan-SUBJ

Luchkina & Cole present an analysis of prominence ratings provided by linguistically naïve native speakers of Russian and find that salient acoustic-prosodic, as well as structural cues to discourse-prominence are used jointly in judgments of the prominence status of a word. The prominence ratings reported in [18] are based on the production of one (model) speaker, which raises the question of whether utterances produced by other Russian speakers would yield similar perceptual ratings.

In this paper we extend our earlier work and explore how prosody and word order function independently and in combination to mark IS in a free word order language like Russian. The proposed experimental design seeks to determine (1) whether in discourse, the sentential position of a word affects its perceived prominence; (2) and whether cross-application of prosodic and structural cues translates into a yet greater degree of discourse prominence.

We test the hypothesis that in free word order discourse, an ex-situ sentential position acts as an independent cue to prominence, and if so, whether prominence may be further reinforced with acoustic-prosodic features associated with such position. To this end, we analyze word order and IS properties of two authentic Russian narratives for acoustic evidence of prosodic marking in relation to word order and IS. We (1) use the word-level prominence ratings obtained in a perceptual prominence rating task to gauge the native speakers' sensitivity to prosodic and structural means of encoding discourse-prominent information and (2) match the two classes of prominence correlates, structural and acoustic-prosodic, with the perceived prominence scores to test the ability of each of these sources of discourse-relevant information to predict the word IS category in the narratives chosen for this study. We analyze the experimental results with respect to the interaction between word order, prosodic marking, and perceived prominence.

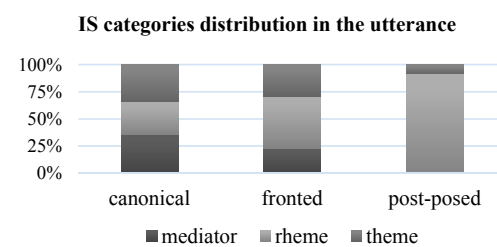
3. Experiment 1: Production task

3.1. Materials

Both discourse samples used in this study come from two published narratives, which display a range of word order and prosodic features. With an average sentence length of 5.2

content words (SD=1.77, 230 content words), approximately 30% of the sentences in the narratives deviate from the canonical SVO order. Adopting the discourse annotation framework introduced in [3], the distribution of the following IS categories¹ was assessed: 'THEME' (word carrying discourse-given information, 2-12 mentions per narrative), 'RHEME' (words carrying discourse-new information, 1st mention in the narrative), 'MEDIATOR' (words carrying inferable information, 1st mention in the narrative), and 'CONNECTOR' (function words). The sentential position of each word in the narratives was marked as in-situ or ex-situ (specifically, 'fronted' or 'post-posed', relative to SVO order).

Figure 1: Observed distribution of IS categories in the analyzed discourses



The observed distribution of IS categories differed by sentential position (Pearson $\chi^2(6)=85.91$, $p<0.001$). Figure 1 illustrates that discourse-novel information is the most 'mobile' information category in the corpus, with more ex-situ locations than any other IS category, and clearly dominates the post-posed (utterance-final) position. Such non-random distribution of discourse information provides evidence that word order variability in the narratives under analysis may be indicative of IS category.

3.2. Acoustic-prosodic features pre-processing & production data overview

For the purposes of the perceived prominence rating task, both narratives were read orally by a female speaker of Russian (henceforth, the model speaker), age 27. Read productions from 14 speakers (ages 21-52, 8 females) were collected for the purposes of extended acoustic-prosodic analysis of the study materials. The acoustic-prosodic measures of f0 and intensity mean, minima, and maxima, f0 range, velocity, and excursion size, as well as vowel duration were taken from every syllable of each IS-coded content word in the corpus. Additionally, distance in milliseconds from the tonal center of gravity of each stressed vowel to the vowel midpoint was measured. All measurements were extracted automatically in Praat [20]. The values of max f0 and max intensity were taken from the center region of the vowel in order minimize the influence of the adjacent segments at the voice onsets and inter-segmental transitions. Each f0 output was transformed to semitone values relative to a fixed value of 100 Hz². Prior to

¹ The IS category of each content word was annotated by one of the authors (TL) and another native Russian speaker. Inter-rater agreement (linearly weighted Kappa) between the annotators, across texts was satisfactory: $\kappa=0.89$, $SE=0.03$, $\alpha=0.05$.

² The semitone scale was chosen to reduce male-female acoustic differences and ensure that +1 standard deviation constituted a perceptually equivalent interval as -1 standard deviation.

being submitted to regression analyses, all acoustic-prosodic measures were centered using mean-centered coding [21].

Extracted acoustic measures were examined as correlates of prosodic prominence, and analyzed for their relationship to the IS and sentential position of a word.

The model speaker's production data were subject to a preliminary analysis to determine how well cross-linguistically attested prosodic correlates of prominence (intensity and f_0 maxima, and duration of the stressed vowel) can be predicted from the word's IS category and position in the utterance. A series of linear regression analyses each featuring one acoustic parameter of interest as the dependent variable revealed that both the ex-situ sentence position and IS category RHEME are reliably associated with greater intensity and f_0 maxima, as well as longer duration of the stressed vowel (see Table1).

Table 1: Significant predictors of vowel intensity, duration, and f_0 , with respect to the carrier word³

(max)intensity	(max) f_0	vowel duration
ex-situ ($t=3.16$, $p<0.005$)	ex-situ, fronted ($t=6.36$, $p<0.001$)	ex-situ ($t=2.57$, $p=0.01$)
RHEME ($t=2.32$, $p=0.02$)	RHEME ($t=2.84$, $p=0.005$)	RHEME ($t=2.94$, $p<0.01$)

Significant patterns of co-variation were observed between some of the acoustic measures: e.g., a significant rise in max f_0 was reliably associated with concurrent increases in intensity ($t=24.77$, $p<0.001$) and duration ($t=5.35$, $p<0.001$). Analysis of the model speaker's production data established that the acoustic parameters of f_0 , intensity, and duration contribute perceptual salience to the words carrying discourse-novel information and/or occurring ex-situ. Before we validate this finding with production data from multiple speakers, we test the psychological reality of the acoustic-prosodic and structural cues to prominence, i.e., gauge their ability to affect reader's or listener's perception of a word as prominent during discourse comprehension.

4. Experiment 2: Prominence rating task

Structural and acoustic-prosodic cues to prominence were determined in reading and auditory comprehension tasks performed by linguistically naïve native speakers of Russian ($N=49$ (reading modality), $N=27$ (auditory modality)). The task included thirty-nine clause-size excerpts from the narratives. Each clause, or target segment, was presented along with the preceding context. The mode of presentation of the target sentence was either written text or audio recording of the model speaker's production. Respondents read the entire portion of the text preceding the target segment, read or listened to the target segment and identified discourse-prominent word(s) in the target segment by associating them with one level of the binary feature "+/- prominent". Following [22], no formal definition of prominence was given. Participants were instructed to mark only those words that 'were the focus of their attention' in the utterance, based on

³ Factors 'speaker' and 'word' (not shown in Table1) were included in the model as random effects.

the preceding context. Any number of content words could be marked as prominent.

4.1. Results

Assessing consistency of the responses: Responses to the prominence rating task were assessed for intra- and inter-rater agreement. The kappa coefficients assessing consistency of the intra-rater rating behavior fall within the range 32.4 – 88.1 (mean=70.4, SE=2.21) and translate into moderate to very high agreement. The inter-rater agreement coefficients translate into fair, though highly significant agreement levels: Fleiss' kappa=0.26 ($p<0.001$) for the written and 0.36 ($p<0.001$) for the auditory modality.

Overall picture of the perceptual prominence ratings: Following [9], each word in the narratives was assigned two discourse prominence scores (one per test modality). Prominence scores were obtained by dividing the total number of times a word was chosen as salient by the total number of participants who responded to the relevant test question. Additionally, a global prominence score was computed for each word to represent its perceived prominence across the presentation modalities. In order to gauge the respondents' sensitivity to word order changes in the narratives, a two-way analysis of variance assessed prominence of the IS categories represented in the corpus under (a) canonical and (b) non-canonical word order. With the global prominence score used as the dependent variable, the ANOVA crossed the factors IS and Word Order (see Table 2). Results revealed no significant main effect of Word Order ($F=0.137$, $p>0.05$), a significant main effect of IS ($F=47.82$, $p<0.001$) and a significant interaction between these two factors ($F=3.36$, $p<0.01$).

Figure 2: Global perceived prominence scores (y-axis) of the 4 IS categories (x-axis) as determined by the combined results of the silent reading and listening prominence rating tasks.

Global perceived prominence scores by IS category & WO

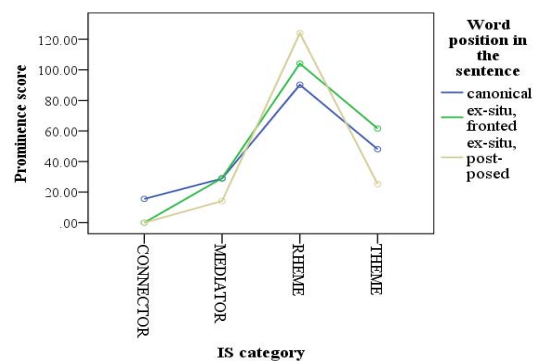


Figure 2 illustrates that IS categories with highest prominence ratings, THEME (discourse-given information, mean prominence score =49.6, SE=6.0) and RHEME (discourse-novel information, mean prominence score =106.0, SE=3.5), demonstrate a meaningful dissociation in the prominence scores as a function of the word position in the sentence. Pairwise comparisons reveal that both *in-situ* (canonical) and *fronted* words carrying *discourse-given* information are viewed as significantly more discourse-prominent than those post-posed relative to their canonical position (mean

difference=22.83, $p=0.05$ for in-situ words, and mean difference=36.4, $p<0.05$ for fronted words). The dissociation goes the opposite way for the ex-situ words carrying *discourse-new* information (RHEME): *post-posed* discourse-novel words receive significantly higher discourse prominence scores than those appearing the in-situ or fronted, relative to their canonical position (mean difference=33.7, $p<0.001$ for in-situ words, and mean difference= 19.9, $p=0.05$ for fronted words).

Correlates of perceived prominence: Modality-specific prominence scores were modelled with linear regression analyses. Predictors for auditory prominence perception were extracted from the recording of the model speaker's reading performance and included the acoustic-prosodic measures of standardized f0 and intensity maxima, f0 range and excursion size, and vowel duration taken from the stressed syllable of each IS-coded content word.

In silent reading of the Russian corpus, words located in ex-situ position, specifically, *post-posed* ($t=7.17$, $p<0.001$) or *fronted* ($t=5.78$, $p<0.001$), relative to the canonical position, as well as words carrying *discourse-novel* information ($t=2.69$, $p<0.01$) were associated with higher prominence scores.

In the auditory modality, these factors were complemented with acoustic predictors duration ($t=2.23$, $p<0.05$), intensity ($t=2.27$, $p=0.05$), and f0 range ($t=2.78$, $p<0.01$).

4.2. Introducing production data from multiple speakers

While results of the auditory prominence rating task present compelling evidence for structural and prosodic prominence being utilized by the listeners during discourse comprehension, they hinge on the reading performance of the model speaker. To test whether the perceived prominence scores from the auditory modality are consistent with the IS properties of the narratives encoded via word order and multi-speaker prosody, acoustic-prosodic measurements of the model speaker's reading performance were augmented with those from 14 speakers (who did not participate in the prominence rating task). A multinomial logistic mixed effects regression was fit to the data to model the IS category of a word (*discourse-given* or *THEME*, *discourse-novel* or *RHEME*, *inferable* or *MEDIATOR*) using its auditory modality prominence score, sentence position, and the acoustic measures for that word extracted from the productions of the 15 speakers⁴. Results of the logistic regression analysis confirm that structural and prosodic correlates of discourse prominence successfully predict the IS category of content words. Specifically, *discourse-novel* information, which is the IS category that received the highest perceived prominence scores ($z=6.69$, $p<0.001$), bears a reliable association with ex-situ sentential positions ($z=1.77$, $p=0.08$ for fronted RHEMES; $z=1.97$, $p<0.5$ for *post-posed* RHEMES) and distinctive acoustic-prosodic realization. The latter is supported by a number of parameters from oral productions by 15 Russian speakers, including f0 range ($z=1.99$, $p<0.05$) and maxima ($z=3.23$, $p=0.001$), lower mean f0 values ($z=-3.18$, $p=0.001$), as well as greater duration ($z=2.10$, $p<0.05$) and mean intensity ($z=2.28$, $p<0.05$) associated with the stressed vowel

⁴ The regression model included two random effects, *segment* and *speaker*. Each level of the variable *segment* corresponds to one unique vowel from which the acoustic-prosodic measurements were extracted.

of RHEMES. *Discourse-inferable* words (MEDIATOR), on the other hand, are characterized by lower perceived prominence scores ($z=-4.84$, $p<0.001$) and smaller duration ($z=-4.30$, $p<0.001$) and mean f0 values ($z=-2.12$, $p<0.05$).

5. Discussion

The goal of this work is to parameterize perceived prominence in a free word order language like Russian and understand which factors guide naive readers' or listeners' perception of a word as prominent in a discourse or narrative. To this end, we offer an empirical test of whether variation in word order, along with the more established acoustic-prosodic ways of marking discourse-prominent information, can serve as a means of encoding the information status of a word and, by doing so, mediate its perceived prominence.

Analyses of the production and perceptual prominence ratings data presented in this work successfully capture the close interrelatedness of the prosodic and structural cues to prominence in the corpus of two published narratives. Results demonstrate that independent of the modality of presentation, words that carry *discourse-novel* information are perceived as more prominent.

In a free word order language such as Russian, information status is encoded via two routes, prosodic and structural. In the auditory modality, listeners treat the acoustic-prosodic realization of a word as a cue to its discourse status. This is evident from the finding that distinctive f0 qualities, greater intensity and vowel duration reliably trigger perception of a word as prominent. About 30% of the utterances in the mini corpus deviate from canonical word order, which means that words carrying novel or given information may occur ex-situ, i.e., be fronted or *post-posed*, relative to the canonical SVO position. Results of the prominence rating task show that apart from the acoustic effects of prosody, an ex-situ position of a word also contributes to its perception as prominent.

Analysis of the syntactic and acoustic-prosodic characteristics of perceived prominence reveals that Russian allows cross-application of different cues to prominence within the same utterance. Ex-situ positions are associated with (1) a higher prominence score and (2) in the auditory modality, distinctive perceptual qualities. Such distinctive acoustic-prosodic realization of non-canonically positioned words may not only cue their relatively high informational load and discourse prominence, but, in a language that exhibits focus fronting and right-edge dislocation for IS purposes [16,25], may also (redundantly) signal that the word is left- or right- dislocated.

6. Conclusion

This study contributes to the understanding of discourse-prominence in a free word order language. Results of the production and perception experiments performed by linguistically naive native speakers of Russian reveal that concurrent application of prosodic and structural cues does not preclude either cue class from being perceived as a signal to information that calls for special attention in discourse. Further work is necessary to determine whether cross-application of prominence cues is characteristic of all vs. select categories of discourse-prominent information and whether its effect is additive, i.e., leading to a word being associated with a yet greater degree of perceived prominence.

7. References

- [1] Watson, D. G. (2010). The many roads to prominence: Understanding emphasis in conversation. In B. Ross (Ed.), *The Psychology of Learning and Motivation*, 52, 163-183. Elsevier.
- [2] Chafe, W.L. (1976). Givenness, contrastiveness, definiteness, subjects, topics, and point of view. In Li, C. (ed.), *Subject and Topic*. New York: Ac. Press, 25-55.
- [3] Arnold, J.E., Losongco, A., Wasow, T., & Ginstrom, R. (2000). Heaviness vs. newness: The effects of structural complexity and discourse status on constituent ordering. *Language*, 76, 28-55.
- [4] Calhoun, S., M. Nissim, M. Steedman, and J.M. Brenier. (2005). A framework for annotating information structure in discourse: Pie in the Sky. Proceedings of the workshop, ACL, 45-52.
- [5] Katz, J., Selkirk, E. (2011). Contrastive focus vs. discourse-new: Evidence from phonetic prominence in English. *Language*, 87, 771-816.
- [6] Morgan, J. L., Meier, R. P., & Newport, E. L. (1987). Structural packaging in the input to language learning: Contributions of prosodic and morphological marking of phrases to the acquisition of language. *Cognitive Psychology*, 19, 498-550.
- [7] Stolterfoht, B., Friederici, A., Alter, K., and A. Steube (2007). Processing focus structure and implicit prosody during reading: Differential ERP effects. *Cognition* 104, 565-590.
- [8] Watson, D. G., Arnold, J. E., & Tanenhouse, M. K. (2008). Tic Tac TOE: Effectiveness of predictability and importance on acoustic prominence in language production. *Cognition* 106(3), 1548-1557.
- [9] Mo, Y., Cole, J., & Lee, E. (2008). Native listeners? Prominence and boundary perception. *Speech Prosody* 2008.
- [10] Kaland, C., Krahmer, E. & Swerts, M. (2011). Contrastive intonation: speaker- or listener-driven? In W.S. Lee & E. Zee (Eds.), *Proceedings of the 17th International Conference of Phonetic Sciences*, 1006-1009. Hong Kong: City University of Hong Kong.
- [11] Donati, C. & N. Nespors (2003). From Focus to Syntax. *Lingua*, 113-11, 1119-42.
- [12] Neeleman, A., Reinhart, T. (1998). Scrambling and the PF interface. In: Butt, M., Geuder, W. (Eds.). *The Projection of Arguments: Lexical and Compositional Factors*. CSLI Publications, Stanford, CA, 309-353.
- [13] Calhoun, S. (2010). The Centrality of Metrical Structure in Signaling Information Structure: A Probabilistic Perspective, *Language*, 86(1), 1-42.
- [14] Skopeteas, S. & Fanselow, G. (2010). Focus in Georgian and the expression of contrast. *Lingua*, 120, 1370-1391.
- [15] Sekerina, I. A. (1999). The Scrambling Complexity Hypothesis and Processing of Split Scrambling Constructions in Russian. *Journal of Slavic Linguistics*, 7.2, 218-265.
- [16] Slioussar, N. (2011a). Processing of a free word order language: The Role of Syntax and Context. *Journal of Psycholinguistic Research*, 40:291-306.
- [17] Arvaniti, A. & Adamou, E. (2011). Focus expression in Romani. In *Proceedings of the 28th West Coast conference on Formal Linguistics*, Somerville, MA: Cascadia Proceedings Project.
- [18] Luchkina, T. & Cole, J. S. (2013). Routes to Prominence in Free Word Order Language Discourse, in Mertens, P. & A.C. Simon (eds), *Proceedings of the Prosody-Discourse Interface Conference 2013 (IDP-2013)*. Leuven, September 11-13, 2013, pp. 13-19.
- [19] Boersma, P. & Weenink, D. (2013). Praat: doing phonetics by computer [Computer program]. Version 5.3.60.
- [21] Mitchell, M. (2012). *Interpreting and visualizing regression models using Stata*. Stata Press, College Station, TX.
- [22] Cole, J., Mo, Y., & Hasegawa-Johnson, M. (2011). Signal-based and expectation-based factors in the perception of prosodic prominence. *Laboratory Phonology*, v.1, 2011, p. 425-452.
- [23] Neeleman, A., Titov, E. (2009). Focus, contrast, and stress in Russian. *Linguistic Inquiry* 40, 514-524.