PROSODIC AND STRUCTURAL VARIABILITY IN FREE WORD ORDER LANGUAGE DISCOURSE

BY

TATIANA LUCHKINA

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Linguistics
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2016

Urbana, Illinois

Doctoral Committee:

Professor Jennifer Cole, Chair
Professor José Ignacio Hualde
Associate Professor Tania Ionin
Associate Professor Ryan Shosted

ABSTRACT

Independently of the modality of presentation (written or auditory), human processing of discourse obligatorily involves monitoring relative information prominence which reflects how important information is in discourse, and thereby determines the perceptual impact it makes on the speaker and the listener. In spoken language use, relative information prominence is expressed it by means of morphology, structural organization of information across an utterance, and by prosodic means (Morgan, Meier, & Newport 1987, Stolterfoht, Friederici, Alter, & Steube 2007, Watson 2010). To illustrate, Hindi speakers may use special morphological prominence markers, bound morphemes 'hii' and 'bhii', which attach to words that the listener is likely to identify as prominent (Luchkina, Puri, Jyothi, Cole 2015). In Hungarian and Hindi, speakers place the prominent word in the pre-verbal position in a sentence or phrase, which presents a designated location for prominent (focused) information in these languages (Genzel and Kügler 2010, Féry 2013). In English, it is the utterance-final or, else, the most prosodically prominent word in a sentence or phrase that is likely to be identified as prominent (Ladd, 2008, Watson 2010).

This thesis examines the use of acoustic-prosodic cues and constituent ordering in the expression of relative information prominence and the way it affects perception (as perceived prominence) in Russian, a free word order language, by empirically testing the "dual route" model of expressing prominence in discourse. This model presupposes (1) structural "packaging" of information, evident from the linear ordering of words in an utterance such that words communicating relatively more accessible and therefore less salient information *precede* words communicating less accessible and therefore more salient information, and (2) varying magnitude of acoustic-prosodic parameters in a controlled way such that prominent information

bears greater perceptual salience in speech. Speech production and comprehension experiments described below test whether these routes, structural and acoustic-prosodic, are used independently or together in the encoding of information prominence. Russian is chosen as the test case because it allows but does not require surface reordering of sentential constituents for information structural purposes and exhibits distinctions in prosodic prominence among the constituents of a sentence (Sekerina 2003, Slioussar 2011a, b, Svetozarova 1998).

To examine how prosodic and structural cues are utilized during the off-line and the online processing of discourse in Russian, the following research objectives are pursued. In Study 1 (see Chapter 2 of the present version), the distribution of structural and acoustic-prosodic variability in read discourse is examined in association with two well-known prominence scales: distinctions in the information status of a discourse referent and animacy of a discourse referent (in conjunction with grammatical function of the corresponding lexical word). In Study 2 (see Chapter 3 of the present thesis), relative contribution and perceptual validity of linearization prominence cues and acoustic-prosodic prominence cues is examined using perceived prominence ratings solicited from linguistically-naïve native speakers of Russian. In Study 3 (see Chapter 4 of the present thesis), processing costs associated with these prominence cues are gauged using probe recognition response times obtained during online comprehension of discourse fragments with experimentally controlled variation in word order and acoustic-prosodic expression.

The experimental investigations reported in this thesis advance the scientific understanding of prosodic and structural variability in read discourse in Russian, a language in which about 20% of spoken utterances deviate from the canonical SVO order (Lobanova 2011, Sekerina 2003). Rigorous examination of acoustic-prosodic dynamics and discourse-motivated word order

modification associated with relative information prominence bears relevance for linguists, psychologists, and teachers of Russian as a Second Language interested in the interaction between intonation and word order in a free word order language and its consequences for language production and comprehension. The findings reported in this thesis also bear relevance for work on speech synthesis in free word order languages, such as Russian, where understanding the effects of word order variability on intonation in spoken language use is among the relevant research goals.

ACKNOWLEDGMENTS

I would like to thank my academic advisor, Professor Jennifer Cole, for her infallible support and encouragement and profound contribution to the research reported in this thesis and beyond.

I would like thank Professor Tania Ionin for working with me from day one in the program, introducing me to experimental work and data analysis and for serving as an example of outstanding work ethics and intellectual integrity.

I would like to thank Professor Ryan Shosted for help with representing production data reported in Chapter 4 of this thesis and serving on my dissertation committee.

I would like to thank Professor José Hualde for serving on my dissertation committee.

I would like to thank Gökhan Torna for introducing me to STATA and providing steady support throughout the years that I have been living away from my family.

Thank You!

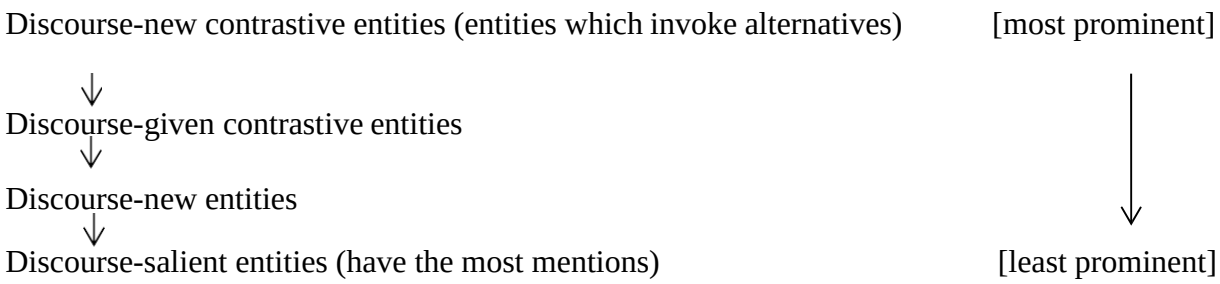
TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

Prosody–word order interface in the expression of discourse-prominent information

Relative information prominence and its impact on how information is perceived in discourse have been offered a variety of interpretations in the linguistic and psychological literature (see Adli 2011, Watson 2010 for review). Numerous studies of relative information prominence (Chafe 1976, Calhoun, Nissim, Steedman, & Brenier 2005) often emphasize distinctions in information accessibility as one of the determinants of its importance or prominence. Most researchers distinguish between discourse entities that are most accessible (or given, active), least accessible (or novel, inactive) and inferable (semi-active). Traditionally, discourse-novel entities are seen as more prominent than the given ones. This is reflected in the tendency to treat new information and contrastively prominent information (a contrastive interpretation is achieved via invoking an alternative set of entities which contrast a given discourse entity), as two information categories which have a relatively higher degree of discourse prominence, particularly according to the work following the Information Theoretic tradition (see, for example, Calhoun 1995, 2010 and Baumann and Riester 2012, 2013). Distinctions in the information accessibility or information status of discourse entities inform the following hierarchy of prominence proposed Wagner, Breen, Flemming, Shattuck-Hufnagel, & Gibson (2010) (based on speech data from American English):



Reflecting on which factors contribute to, and manifest relative information prominence, Cole, Mo, & Hasegawa-Johnson (2011) list lexical frequency and predictability of a word, as well as speaker-dependent variables, such as speech rate. Prominence conceptualizations such as these reveal that a variety of factors including information status, lexical choices, and pragmatic phenomena, including focus or emphasis, convey relative information prominence status and determine the subsequent perceptual impact that information is entitled to in discourse. Consistent with these lines of thought, in this thesis, prominence is understood to be a relatively greater perceptual saliency of information expressed by prosodic means at the phonological level, domain of intonational phrase (IP), and linearization means at the morpho-syntactic level, domain of a sentential clause. During communication, prominence is utilized to signal words that are in focus of the interlocutors' attention.

A variety of syntactic, prosodic, and morphological factors have been cross-linguistically attested to manifest relative information prominence in discourse (Morgan, Newport, & Meier 1987). The following discussion selectively focuses on word order, a morpho-syntactic tool which gives rise to structural variability, and acoustic-prosodic means, which give rise to perceptible variability in the acoustic-prosodic parameters in spoken or read discourse. Both these sources of variability may affect the likelihood that a word is perceived as prominent. Following Sneed (2004), in this work, the term '*structural prominence*' refers to special ordering of sentence constituents in a clause which results in one constituent being rendered as more prominent than the rest. The second

type of variability, acoustic-prosodic, gives rise to *'prosodic prominence'*. As the term implies, it involves changes in the prosodic parameters which render a word audibly augmented or reduced and thereby affect the likelihood that it is perceived as prominent. Next, we review how each of these sources of variability in discourse is involved in cuing relative information prominence.

Prosodic correlates of prominence

Prominence signaled via prosodic means is a psychologically real property of discourse production and perception, as supported by much empirical work (see, e.g., Watson et al. 2008, Mo, Cole, Lee 2008, Kaland, Krahmer, Swerts 2011). Prosodic encoding specific to discourse-prominent information involves perceptually salient changes in the voice quality of the speaker, which include duration and intensity assigned to various discourse segments, as well as changes in the fundamental frequency, or pitch. For example, cross-linguistically, words carrying novel information, as well as discourse-prominent words, are known to be associated with a perceptually salient rise in pitch (pitch accent) accompanied by greater duration of the stressed vowel (Breen, Fedorenko, Wagner, & Gibson 2010).

In English, a relatively fixed word order language in which prosodic prominence is the major way of communicating relative information prominence, phonological prosodic structure is phonetically realized through a variety of acoustic properties that distinguish articulation of the prominent word compared to the articulation of the surrounding words that are non-prominent. In a survey of more than twenty acoustic properties of the speech signal that can potentially correlate with different categories of information structure in English, Breen and colleagues conducted a stepwise discrimination function analysis and found that greater intensity, longer duration, and higher mean and maximum f0 are all reliable predictors of the location of focused words. More careful articulation of the prominent word enhanced with the higher mean values of

f0, intensity, and segment duration have been also recorded in analyses of phrasal prominence in English by Ladd (2008), Kochanski, Grabe, Coleman, & Rosner (2005), & Beckman (1986).

Special prosodic expression of prominent information is also in use in languages which, unlike English, demonstrate word order variability. To illustrate, f0 and duration are important acoustic-prosodic correlates of prominence in European Portuguese, Greek, Dutch, Italian, and Romanian (Frota 2002, Swerts 2007, Swerts, Krahmer, & Avesani 2002). Unlike in fixed word order languages, prominence is additionally associated with dedicated clausal positions in languages displaying word order variability. This is evident from the fact that words which introduce discourse-new or contrastively focused information often 'favor' specific positions in a sentence or phrase, such as preverbal, clause-final or clause-initial. Despite these linearization preferences in the free word order languages, exceptionally few are known which *do not* manifest prominence prosodically (e.g., Wolof (Rialland 2001).

Treating perceptible changes in acoustic-prosodic parameters as discourse-motivated, Bolinger (1972), Selkirk (1996), and Calhoun (2010) argue that prosodic prominence is usually assigned to a word carrying information that is unexpected or new, salient, and/or focused. Aylett & Turk (2004), Watson (2008), and Cole, Mo, & Hasegawa-Johnson (2011) argue that the informational load and predictability of a word, as well as its lexical frequency reliably trigger variation in acoustic-prosodic expression as a discourse prominence. Given multiple potential factors which correlate with prominence in discourse, Watson (2008) emphasizes that it should be thought of as a 'multi-source phenomenon'. Similarly, rather associating prosodic expression of prominence with any given acoustic-prosodic correlate, such as word duration or intensity, it should be viewed as a complex interaction of speaker-based and hearer-based components that reflect the cognitive processes of discourse generation and perception. Another important assumption about the nature

of (prosodic) prominence made in work by Watson and Mahrt, Cole, and Fleck, and Hasegawa-Johnson (2012) is that it should not be conceived as a strictly categorical phenomenon but as a continuous measure which varies in relation to relative information prominence and accessibility in discourse. Treating information accessibility as gradient or continuous is also advocated in the work by Slioussar (2011b), who interprets accessibility as 'focus of attention […] converted to discrete categories on the way from intention to articulation' (p.6). According to Slioussar, the least or the most accessible entity receives distinctive prosodic expression via prosodic augmentation or reduction. Slioussar highlights that the same effects can be relayed by structural means, e.g., via topicalization of the more accessible information and post-posing of the least accessible information, implemented via monitoring the linear succession of distinct information categories in an utterance. Consistent with this view, let us now discuss how relative information prominence can be relayed by structural means, or strategic positioning of the discourse-prominent word in relation to the surrounding context.

Structural correlates of prominence

Unlike in rigidly ordered languages, where the syntactic function of a word is determined by its position in a sentence or phrase, morphological case is what determines the syntactic function of a word in free word order languages, such as Finnish or Turkish. In these languages, rich morphological paradigms is what makes the word order feature salient and interpretable on its own and brings about a variety of syntactic, semantic, morphological and phonological changes that distinguish non-canonically ordered utterances from their canonically-ordered counterparts (Horwath, 2010, Arvaniti & Adamou 2011).

In a language with rigid word order, little variability in constituent ordering is available for the speaker to control how soon a word carrying prominent information occurs in a sentence or

phrase. Under free word order, ordering preferences appear to be motivated by information structure and relative information accessibility, rather than by grammatical function of a sentence constituent. This is possible because languages with free word order allow for various permutations of sentential constituents, as shown in examples (3-5) below, and deploy this flexibility in the surface constituent ordering to maintain an optimal distribution of novel and given information across an utterance (Sekerina 1999, Slioussar 2013). Such preferred distribution, cross-linguistically, leads highly accessible given information, or THEME, to precede less accessible or novel information, or RHEME and facilitates processing and subsequent recall of information. Relevant experimental evidence comes from an ERP study by Pyykkönen, Drenhaus, & Crocker (2011) who found that in German, listeners have expectations during comprehension of discourse to have more salient discourse entities mentioned prior to less salient entities, and to have discourse-given entities mentioned prior to the discourse-new entities[1]. Pyykkönen and colleagues conclude that in German ditransitive verb constructions, speakers prefer to use a previously mentioned entity as the first object of a verb and a novel entity as its second object. Reversing the order of the entities is perceived as infelicitous, even though both constituent orders are grammatical in the language. That ordering of given and novel information in discourse affects the ordering of sentence constituents was also reported for Finnish, a highly free word order language, by Kaiser and Trueswell (2004). Kaiser and Trueswell found that in Finnish, canonically an SVO language, starting a sentence with an object is only felicitous when the object is associated with given or established information in discourse and the subject communicates information which has not been previously mentioned.

[1] Similar expectations have been attested for English (Clifton & Frazier 2004, Arnold, Losongco, Wasow, & Ginstrom 2000, Bock & Warren 1985).

Findings of Pyykkönen et al. and Kaiser and Trueswell, as well as other related work summarized in Chapters 2 and 3, confirm that by placing a word into a designated position that accords with its information status is an effective way of enhancing comprehension of discourse and facilitating subsequent recall of information.

Relatively few languages are known which *require* surface rearrangement of sentence constituents to express focus, emphasis, and relative information prominence. According to Szendrői (2003) and Sneed (2004), Hungarian is one such language, as it requires that the prominent (focused) constituent occurs pre-verbally and is leftmost in the prosodic phrase, in order to be realized as prosodically prominent.

Obligatory structural reorganization has been also documented for Wolof (Rialland & Robert, 2001) and Chickasaw (Gordon, 2007), where the use of focus particles triggers a set of required syntactic changes at the utterance level. Arguably, narrow and/or contrastive focus *in situ* accompanied with stress shift is infelicitous in Spanish, where discourse-prominence is strongly associated with sentence-final position, which is also the locus of the main sentence stress (Zubizarreta 1998, Hoot 2012).

To summarize so far, relative information prominence may be expressed by prosodic and structural means. During spoken language use, *prosodic* expression of prominence may be observed independently of constituent linearization in an utterance (Hirst & Di Cristo 1998), which is rarely obligatory.

Simultaneous availability of prosodic and structural cues to prominence

In this thesis, the scenario in which *prosodic and structural* cues to prominence are simultaneously available in a language and may possibly be used concurrently is of special interest. To satisfy these conditions, a language under investigation should (a) allow but not

require surface reordering of sentence constituents for information structural purposes and (b) have an established use of prosodic cues to relative information prominence. If both of these conditions are satisfied, it is relatively uncommon that a change in word order is required in order for a word to receive special prosodic expression rendering it prominent. As a result, even under free word order, prosodic prominence can be expressed in-situ. This observation led some researchers to conclude that when prosodic and structural variability are simultaneously available to a language user, expressing relative information prominence is achieved via one of these routes (Swerts, Krahmer, Avesani 2002; Donati and Nespor 2003).

Based on a comparative study of Italian and Turkish (relatively free word order languages), and English (a fixed word order language), Donati and Nespor (2003) propose that languages with rigid word order deploy prosodic marking of prominent information at different locations in the sentence, while languages with flexible word order exhibit less variation in the location of prosodic prominence, and vary the location of sentence constituents instead. This model, with two alternative routes for expressing prominence, predicts that it will be relatively uncommon that a language uses both word order *and* prosodic marking simultaneously to manifest prominence. With respect to this, Donati and Nespor suggested to categorize languages into *prominence dislocating* (e.g., English), i.e., those which utilize prosodic cues to express prominence, and *constituent dislocating* (e.g., Turkish or Italian), i.e., those which primarily use structural means, i.e., constituent ordering in a sentence or phrase.

The phenomenon of metrical reversal or stress shift (Neeleman and Reinhart 1998) illustrated by Calhoun (2010) shows how prominence displacement (or, using Donati and Nespor's terminology, prominence dislocation) works in English. Rightmost accent placement is the default location of the prominent constituent in English, as shown in (1.1) below.

(1.1) Joel bought a green PORSCHE.[2] (1.2) Joel bought a GREEN porsche.

(Examples from Calhoun 2010, p.15).

Rendering a different constituent prominent requires the speaker to modify the prosodic realization of the utterance by displacing the accent, i.e., shifting its location leftwards, as shown in (1.2). In (1.2), by displacing the prominence, the speaker acoustically singles out the constituent that does not belong to the neutral focus set of the utterance resulting in a different interpretation of the sentence by the hearer. Examples (1.1) and (1.2) illustrate how the mechanism of deaccenting moves accents leftwards in English to signal that the constituent receiving prominence is not rightmost in the prosodic phrase. Ladd (2008) emphasizes that such prominence marking mechanism is attested in languages which manipulate the location of prosodic prominence in a sentence or phrase: in these languages, deaccenting of a sentence constituent means making another constituent prosodically prominent, via accenting.

Examples (1.3)-(1.5) below illustrate a different prominence marking strategy, which Donati and Nespor (2003) refer to as 'constituent dislocation'. Example (1.3) demonstrates that in languages with relaxed constituent ordering, such as Italian, instead of altering the placement of prosodic prominence in an utterance, speakers may choose to move the prominent word to a designated location and thereby render that word prominent. This surface reordering of sentence constituents may be prosodically motivated, which means that a word undergoes overt movement in order to be associated with the main phrasal prominence. The location of such 'default' prominence position differs depending on the head-complement parameter and canonical word order of a given language (Ladd 2008, Selkirk 1984). Selkirk (1984) observes that head-initial languages, such as Italian routinely place the prominent constituent into the rightmost position in the

[2] The words in CAPITAL letters are prosodically prominent.

intonation phrase, as shown in (1.3), whereas head-final languages (e.g., Hindi or Turkish) reserve the rightmost spot for the verb and place the prominent constituent pre-verbally, as shown in the Turkish examples (1.4) and (1.5). Constituent dislocation to a pre-verbal position in (1.4) alters the interpretation of the sentence with special emphasis given to the word which word occupies the pre-verbal position.

Italian (head-initial, canonical word order: SVO):

(1.3) E`arrivato Mario.
arrived Mario
'Mario arrived'

Turkish (head-final, canonical word order: SOV):

(1.4) Ahmet visneli keki Anya-ya verdi.
Ahmet-NOM cherry cake-ACC Anya-DAT gave
'Ahmet gave Anya a cherry cake'.

(1.5) Anya-ya visneli keki Ahmet verdi.
Anya-DAT cherry cake-ACC Ahmet -NOM gave
'Ahmet gave Anya a cherry cake'.

Constituent dislocation is a known strategy for expressing contrastively prominent information, or contrastive focus. To illustrate, Neeleman and Titov (2009), Slioussar (2011b), Skopeteas and Fanselow (2010) examined derivation of ex-situ contrastive foci in Russian and Georgian, languages which are known to optionally front contrastively focused constituents, by positing a two-step overt movement, as follows. To render a word focused, a constituent moves to the intonation phrase-final position, or receives prominence locally if positioned IP-finally (1-step derivation). Further, to render a word contrastively focused, a second derivation step is optionally available, wherein the focused constituent moves to the left edge of the intonation phrase.[3] As no

[3]But see a study by Bartels and Kingston (1994) who failed to find any empirical validation of

apparent syntactic reason for this second movement has been proposed (Neeleman & Titov 2009), it is considered to reflect individual speaker choices (Skopeteas & Fanselow 2010) or serve rhetorical purposes (Slioussar 2011b).

Some empirical challenges

This work draws attention to the growing body of empirical evidence that word order variability often presents an optional resource for expressing prominence, available to speakers of free word order languages *along with* acoustic-prosodic means (Skopeteas and Fanselow 2011, Sekerina 1999, Slioussar 2011a, Arvaniti and Adamou 2011, among others). Although expressing prominence via combined structural and prosodic mechanisms presents a challenge to Donati and Nespor's proposal summarized above, numerous languages which deploy surface constituent movement for information structural purposes, including Greek, Russian, Georgian, Finnish, and Hindi, are also known to use prosodic means to mark distinction in the information status and relative information prominence. Expressing (new information) foci in-situ may in fact be preferred over movement in some highly free word order languages such as Georgian (see Skopeteas and Fanselow 2011). This clearly contradicts the phonological/syntactic dichotomy advocated by Donati & Nespor, who maintain that their account of (prosodic) prominence and constituent dislocation is particularly relevant for how non-contrastive new information foci are expressed cross-linguistically. A further challenge to viewing prosodic and structural prominence as complementary comes from the steadily accruing evidence for cross-application of these distinct cues.

Production evidence. A number of studies which investigate how focused information is expressed in free

contrastiveness being encoded any differently than other kinds of discourse-prominent information.

word order languages demonstrate that combined use of prosodic structural cues to prominence is attested, cross-linguistically. For example, Botinis (1998), Skopeteas, Féry, and Asatiani (2009), Skopeteas and Fanselow (2011) report that in Greek, Russian, and Georgian, highly free word order languages, distinctive prosodic expression of discourse-prominent information is observed independently of its position in an utterance. More specifically, in a study of word order and intonation in Georgian, Skopeteas et al. (2009) worked with semi-elicited question-answer pairs in which the answer component displayed different degrees of deviation from Georgian canonical order, SOV.

Figure 1.1. *f0 contours of the utterances in (1.6) (top) and (1.7) (bottom). Reprinted from Skopeteas and Skopeteas et al. 2009, pp. 106-107.*

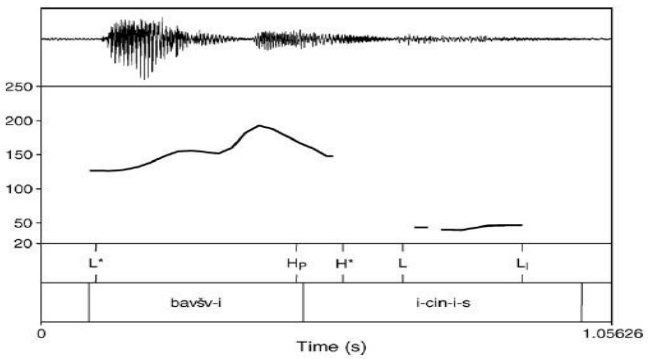


Fig. 1. All-new realization.

(1.6) {What is happening?}
[[bavsˇv-i]P [i-cin-i-s]P]I
child-NOM PV-laugh-PRS-S.3.SG
'The child is laughing.'
<all new information>

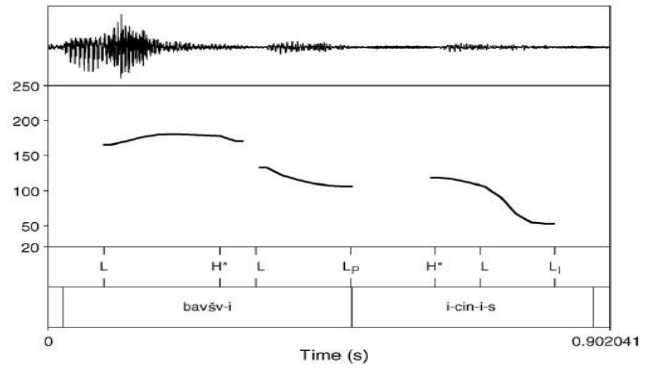


Fig. 2. Narrow focus realization.

(1.7) {Who is laughing?}
[[bavsˇv-i]P [i-cin-i-s]P]I
child-NOM PV-laugh-PRS-S.3.SG
'The child is laughing.'
<new information: CHILD>

Production data presented by Skopeteas et al. contain examples of Georgian sentences in which prominence displacement, rather than constituent displacement is used to signal the discourse prominent constituent, as shown in (1.6) and (1.7), even though Georgian is characterized as a language of 'extreme word order freedom' (p.102). Additionally, Skopeteas and colleagues found

that each word order configuration in Georgian is licensed by a special prosodic realization that makes deviations from the canonical word order perceptually distinct. In a follow up study, Skopeteas and Fanselow (2011) reiterate that structural and prosodic cues to discourse-novel information may be used in concert in Georgian. The authors conclude that their data do not allow for a coherent account of when or why a change in word order gets reinforced with acoustic-prosodic augmentation. The authors tentatively propose that while using the acoustic-prosodic route may present a more economical way of expressing focus in Georgian, cross-application of constituent dislocation and acoustic-prosodic augmentation may reflect a personal preference of the speaker. That prosodic and structural prominence-lending cues can be utilized concurrently has also been reported for the variety of Romani spoken in Komotini by Arvaniti and Adamou (2011). Arvaniti and Adamou observed that the preferred way to mark focus and topicalization in Romani is via a combination of prosodic and morpho-syntactic strategies which include a change in word order (VO→OV) and, for expressing focus, deaccenting of the material to the right of the focused word. Arvaniti and Adamou (2011) second Skopeteas and Fanselow (2011) by saying that the concurrent use of acoustic-prosodic and structural cues to focus is not a necessity in Komotini Romani, but, rather, a matter of speaker preference.The findings reviewed here and in the subsequent chapters posit a problem for the account proposed by Donati and Nespor (2003) in that they demonstrate the following. First, a language may have the necessary resources to optionally or obligatorily deploy acoustic-prosodic and structural cues to signal relative information prominence. Second, speakers may maintain a preference as to which strategy, prominence displacement or constituent displacement, or both, they use to manifest prominence.

Processing evidence. Prosodic augmentation via pitch accenting is a potent mechanism of

information retention in the memory. Accented information is not only retained in the memory with more facility, but is recalled more accurately even 24 hours after it has been presented (Fraundorf, Watson, and Benjamin 2010). While cognitive benefits of *structurally* prominent ex-situ information have not been experimentally assessed, psycholinguistic investigations have revealed that processing of non-canonically-ordered utterances such as (1.3) and (1.5) is highly context-constrained and resource-intensive (Erdocia, Laka, Mestres-Missé, and Rodriguez-Fornells 2009, Kaan 2001, Sekerina 2003, Kaiser and Trueswell 2004). More specifically, Sekerina (2003) argues that in free word order languages, including Russian, processing non-canonically ordered utterances is more computationally costly, as it involves recovering or reactivating the traces of the extracted arguments at their canonical (syntactically determined) sites. Word order and acoustic-prosodic variability in Russian are what motivates the choice of this language for the studies reported in this thesis. It is these characteristics of the Russian language that are discussed next.

Word order variability in Russian

Russian is traditionally characterized as a non-configurational language in which the ordering of sentence constituents does not necessarily maintain a direct link between the structural position of a sentence constituent and its grammatical function. Russian is often characterized as a 'scrambling' language (Baylin 2002), meaning that it permits sentence constituents to occur in different orders without changing the truth-conditional meaning of a sentence or phrase. Scrambling, therefore, refers to the surface syntactic movement which is motivated by information structural factors (van Gelderen 2003) and is non-hierarchical in nature. Consequently, and similarly to other non-confirgurational languages, word order in Russian is reflective of the information structural relations in a sentence or phrase (Kallestinova 2007, Slioussar 2011b). The

six possible orderings of the constituents in Russian are SVO (canonical), OVS, SOV, OSV, VSO, and VOS. As a result, a simple transitive sentence like 'Ivan is cooking pizza' can be expressed in six different ways, corresponding to six possible word orders, as shown in (1.8) below.

(1.8) a. Ivan gotovit pizzu. b. Pizzu gotovit Ivan. c. Pizzu Ivan gotovit.
d. Gotovit Ivan pizzu. e. Ivan pizzu gotovit. f. Gotovit pizzu Ivan.

In Russian, as in other free word order languages, information structure plays a role in determining the surface constituent order through interaction with the intonation the and morphological system; hence, prosodic and morpho-syntactic properties of sentence constituents are important in determining their information-structural import. In Russian, grammatical function of a sentence argument is typically identified through a morphological (case) marker. The Russian case system is diverse and comprised of the following six cases: Nominative, Genitive, Dative, Accusative, Instrumental, and Prepositional.

As shown in the example 1.8., the subject noun *Ivan* bears Nominative case, and the object noun *pizzu* bears Accusative case. By default, in Russian, subjects are sentence-initial (unmarked) topics and objects- sentence-final (unmarked) foci. By virtue of having an explicit morphological case marker, both these arguments can appear sentence-initially and sentence-finally, as long as the chosen linearization is deemed contextually appropriate. Highlighting the importance of morphological case for word order variability in Russian, Titov (2007) argues that it acts as a formal license for scrambling in Russian, whereby a (surface) reordering of sentence constituents is licensed iff "the grammatical function of at least one of these arguments is identified by a morphological marker it carries" (p. 27). Titov (2007) further notes that despite the well-developed case morphology, cases of isomorphism are not uncommon. For select noun classes (esp. feminine nouns of the 3rd declension type, e.g., 'mat''(Rus. 'mother'), 'doch'' (Rus. 'daughter')), the nominative and the accusative noun morphology are isomorphic, resulting in unavailability of free

constituent order permutations for such sentences. Apparently, structural identification of grammatical function takes places in Russian in cases in which morphological case is not revealing. In the majority of cases, however, the rich inflectional morphology of Russian means that the grammatical function of the sentence argument, as well as the argument structure of the entire clause is transparently signaled by means of the case affixes regardless of where in the sentence an argument surfaces. Rich morphological paradigm of the Russian language allows for a transparent mapping between the surface syntax and the information structure when the canonical SVO order does not endorse one. As a result, surface syntax of Russian provides a rather transparent mapping between the spoken form and discourse information structure.

When presented out of context, all the sentences in (1.8) express the same basic idea and are truth-conditionally equivalent, but once presented in discourse, as shown in (1.9), they are no longer equivalent in terms of their information structure: while (a) displays the default or canonical SVO order and is the least context-dependent configuration, the remaining sentences require a particular distribution of information structure (IS) categories in the preceding context to be regarded as felicitous. Consider the following example from Russian, in which the sentence in (1.9) can be continued as in (a) or (b):

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

a. Ivan gotovit pizzu.
Ivan-SUBJ cooks pizza-OBJ

b. Pizzu gotovit Ivan.
pizza-OBJ cooks Ivan-SUBJ

Both the canonical SVO order (a) and the non-canonical OVS order (b) are possible continuations for the sentence in (1.9). Yet, to be regarded as felicitous, these continuations must be used under different discourse conditions. In the context provided in (1.9), the word *Ivan*, critical to our understanding of who is doing the cooking, occurs sentence-finally, where it is *structurally*

prominent (as in b), or else, if occurring pre-verbally as in (a), is *prosodically-prominent* and perceptually stands out. Example (1.9) illustrates how word order variability in Russian may be used to express information status and relative information prominence.

While truth-conditionally, continuations offered in 1.9 appear identical, it is the non-truth-conditional aspects of meaning related to the information status and pragmatic interpretation that these utterances differ on. Evidence for the importance of word order in the expression of prominence in Russian comes from Malamud (2000) who observed that independently of the prosodic realization of an utterance, the ordering of sentence constituents has an effect on how they are perceived. Malamud analyzed native speakers' coherence and prominence ratings of discourse segments containing canonical vs. non-canonical sentences in Russian. She found, among other thnigs, that locating the subject post-verbally demotes its perceived prominence.

The surface position of a sentence argument, then, serves as an indicator of its information structural status and relative information prominence in discourse. As Titov (2007) argues, '…a syntactic constituent [in Russian] can be associated with such an information-structural interpretation as a result of mapping principles that relate syntactic structures to information-structural representations.' (p.34).

By default, left peripheral constituents in Russian are interpreted as sentence topics (given information) and right peripheral constituents – as sentence foci (novel information). As in other right-branching languages, new information foci in Russian are removed from the left periphery and are typically located sentence-finally. In line with this observation, Krylova and Khavronina (1984) report that the focus exponent in a sentence is its most embedded argument and receives the major phrasal prominence or the nuclear pitch accent. Similarly, King (1995) and Brun (2001) argue that in Russian, certain information structure categories, such as given and novel, are

reliably associated with designated clausal positions, sentence-initial and sentence-final, respectively.

To summarize, non-emotive sentences in Russian have neutral intonation and a nuclear-pitch accented sentence-final constituent, also associated with discourse-novel information. The preferred or unmarked information structural template for Russian is the one in which the sentence subject is presupposed (given) and realized as a sentence topic, whereas the VP is the information focus of the sentence. Under such a template, all constituents remain in-situ and a perfect match between the syntax and the information structure is maintained. When necessary, the rich case morphology enables alignment with the nuclear pitch-accented sentence-final position to be achieved through surface constituent movement.

As in other free word order languages, Russian exhibits topic fronting, via object shift, and permits deletion of highly accessible topical elements as in sentences with PRO-drop or null subjects. With this regard, Titov (2007) comments that highly topical material in colloquial Russian, freely available from background information, normally occurs in the sentence-initial position or omitted. Following Zubizaretta (1998), topic shift may be regarded as an instance of prosodically-motivated surface constituent movement implemented to free up the sentence-final focus position of non-focused material. Such interpretation of topicalization is further supported by the fact that it is an optional dislocation process. The optionality of topic shift may be explained by the observation that topichood is primarily established through context. (Ertschik-Shir 2006). Contrastively focused information, on the contrary, may surface in various positions or remain in situ. More specifically, contrastive interpretation of a focused constituent in Russian is often signaled by focus fronting or focus preposing (Titov 2007). Focus preposing, in general, renders the reading of the focused constituent identificational, and suggests that the referent of a

contrastively focused constituent presents a member of a restricted set whose identification leads to the exclusion of other set members. Van Valin and LaPolla (1997) argue that the clause-initial position reinforces the focal interpretation of the pre-posed contrastively focused constituent. Similarly, ex-situ post-verbal placement of the sentence subjects is related to its contrastive reading, although the latter has sometimes been characterized as topical or focal.

Prosodic manifestation of prominence in Russian

Earlier research has established that stress and intonation are closely related to word order in Russian and are further affected by how given and novel information is distributed across an utterance (Yokoyama 1986, Sekerina 2003, Jasinskaja 2013, among others). In her overview of the intonation system of the Russian language, Svetozarova (1998) argues that Russian intonation possesses 'great distinctive possibilities' (p.265) which are closely related to the word order freedom and the relative scarcity of the formal devices, such as function words or articles, used to express information structure or communicative utterance types (questions vs. statements vs. exclamations) in other languages.

In Russian, information structural relationships can be signaled through a linear ordering of sentence constituents (structurally), but also through their intonational properties (prosodically). Given that the sentence-final word in Russian, by default, receives a nuclear pitch accent and is expected to introduce novel information into discourse (Neeleman and Titov 2009), Svetozarova proposes that prominence dislocation leading to prosodic augmentation of a *non-final* word in a sentence or phrase is the primary way of expressing what she refers to as 'special stress' in the Russian language. Consistent with Donati and Nespor's (2003) proposal, it is the unexpected location of the main prosodic prominence in an utterance, which triggers the interpretation of prosodically prominent word as perceptually prominent.

To date, relatively few studies have looked at the acoustic-prosodic correlates of relative information prominence in Russian. Kochanov (2010) tested a variety of automatic speech recognition algorithms using a corpus of recorded spontaneous speech in Russian. The predictors of prominence tested in Kochanov's study are categorized as a) 'melodic', or pitch-related, b) dynamic, or related to intensity, and c) temporal, or duration-related, and include maximum, mean, and minimum f0, f0 excursion, (mean) word intensity, word length, and phoneme length. More relevant studies on acoustic-prosodic variability in discourse in Russian, related, specifically, to how linear ordering of sentence constituents, not hierarchical in nature, interacts with prosody, are reviewed in the subsequent chapters.

Research objectives

This work investigates the application of, and the relationship between, simultaneously available prosodic and structural variability in Russian. In what follows, a change in word order and a change in acoustic-prosodic expression are examined as cues to relative information prominence and perceived prominence in read discourse. To this end, structural and acoustic-prosodic correlates of prominence in Russian are analyzed using read original (published) and experimentally designed discourse segments. Perceived prominence ratings supplied by linguistically-naïve native Russian speakers and online behavioral measures (response times) are used to gauge relative cue validity and processing costs associated with a change in word order, as well as a change in the placement of the nuclear pitch accent used as cues to prominent information. Special attention is given to contexts in which acoustic-prosodic variability and a change in word order cross-apply to cue relative information prominence in discourse.

The following research goals are addressed in the subsequent studies:

- Determine correlates of relative information prominence in read discourse in Russian

(Study 1, Chapter 2 of the present thesis);

- Examine word order and acoustic-prosodic variability as cues to perceived prominence in read discourse in Russian (Study 2, Chapter 3 of the present thesis);

- Evaluate the relative accessibility and cue validity with relation to prosodic and structural cues in read discourse in Russian (Study 3, Chapter 4 of the present thesis).

CHAPTER 2

RELATIVE INFORMATION PROMINENCE IN READ DISCOUSE IN RUSSIAN

(Paper 1)

Abstract

This study examines prosody in read productions of two published narratives by 15 Russian speakers. Two distinct sources of variation in acoustic-prosodic expression are considered: structural and referent-based. Structural effects refer to the particular linearization of words in a sentence or phrase. Referent-based effects relate to the semantic and pragmatic characteristics of the discourse referent of a word, and to grammatical roles that are partially dependent on referent characteristics. Here we examine referent animacy and the related grammatical function of subjecthood, and the relative accessibility or information status of a word. We evaluate variation in acoustic-prosodic measures of mean intensity, duration, and f0 range observed in relation to these factors. We document patterns of (partial) *prosodic augmentation* and prosodic reduction due to structural and referent-based factors, as evident from increased values of acoustic-prosodic measures. Prosodic augmentation due to structural effects is observed for words positioned ex-situ, independent of their semantic, grammatical or pragmatic features. Prosodic augmentation due to referent-based effects is observed for words that are grammatical subjects with animate referents. Prosodic expression is further affected by referent information status. Discourse-given and discourse-new information show greater prosodic augmentation than inferable information. A closer look at individual speakers' production styles reveals that structural and referent-based variation occur in combination and interact.

Introduction

Languages offer a variety of syntactic, prosodic, and morphological tools to encode the relative accessibility and salience of concepts and entities evoked in everyday language use (Morgan, Meier, Newport 1987). In connected speech or discourse, prosodic expression is one such tool (Pierrehumbert and Hirschberg 1990; Ladd 2008; Watson, Arnold, Tanenhaus 2008; Kaland, Krahmer, Swerts 2011). It involves perceptually salient changes in the voice quality of the speaker, in the relative duration and intensity of various discourse segments, as well as changes in pitch. The grammatical category of a word, its position in a sentence or phrase, and its information status (*new* or *given*) have all been found to affect its prosodic expression. To illustrate, grammatical and linearization effects on prosodic measures are demonstrated in Brazilian Portuguese and Hindi, among other languages. In Brazilian Portuguese, there is an increase in duration and f0 range for grammatical subjects compared to grammatical objects, and an increase for utterance-initial subjects compared to non-initial subjects (Antão, Arantes, Cunha Lima 2013). There are similar prosodic effects in Hindi, an SOV language, for words that are situated in non-canonical positions, e.g., with the placement of an object noun in sentence-initial position (Patil, Kentner, Gollard, Kügler, Féry, Vasishth 2008). Prosodic effects due to information status are demonstrated in numerous studies, most extensively for English, a fixed word order language, but also for languages with more flexible word order, such as Greek, Turkish or Hindi (Baltazani and Jun 1999; Baltazani 2003, Işsever 2003; Patil et al. 2008). For example, English words expressing previously unmentioned novel information, as well as contrastively focused or emphasized words, have been reported to have increased pitch, greater duration, and greater intensity compared to words that are not focused or are already given in the discourse context (Xu and Xu 2005; Mo, Cole, Lee 2008; Breen, Fedorenko, Wagner, Gibson 2010; Cole, Mo, Hasegawa-Johnson 2011).

In Hindi, a free word order language, an increase in intensity and f0 maxima, with expanded pitch range, have been found to mark the (contrastively) focused constituents and novel (and also more prominent) information in discourse (Patil et al. 2008; Genzel and Kügler 2010; Luchkina, Puri, Jyothi, Cole 2015).

An accurate characterization of the prosodic encoding of discourse meaning must take into account the effects on prosody of both *structural* factors, i.e., the position of a word in an utterance, and the properties related to the word's referent, such as inherent semantic features and grammatical functions that are (partly) dependent on those features, and information status. For the purposes of this study, we refer to these non-structural properties as *referent-based* properties. We focus, in particular, on (a) referent animacy analyzed in conjunction with the grammatical function of the corresponding referring expression, and (b) referent information status, which relates the word to the preceding discourse. The present study examines these factors and their interaction in conditioning variation in acoustic prosodic measures in Russian. Russian poses an interesting case study because, as in English, the intonational prominence of a word is sensitive to the position of the word in the prosodic phrase, but unlike in English, word order in Russian is quite variable (Yokoyama 1986; Sekerina 2003). Moreover, variation in word order in Russian may be sensitive to referent features, such as animacy and information status (Lobanova 2011; Jasinskaya 2013), further strengthening the potential for interaction between structural and referent-based factors on the prosodic expression of words.

We examine the effects from structural and referent-based factors on Russian prosody through the analysis of read productions of two published narratives by fifteen Russian speakers, asking how variation in the acoustic-prosodic parameters of intensity, duration, and fundamental frequency are affected by 1) the surface ordering (or linearization) of arguments in a sentence or phrase and 2)

the characteristics of discourse referents, including the inherent semantic feature of animacy and grammatical functions that are sensitive to animacy, and referent information status. In the first series of analyses, we examine changes in prosodic expression triggered by a discourse-motivated change in word order. In the following discussion, we determine if readers' prosody reflects referent-based properties. Specifically, we look for prosodic effects due to animacy in conjunction with grammatical function (subjecthood), and prosodic effects of the information status of discourse referents. We present evidence for distinct effects on prosodic expression, structural and referent-based. Our findings also reveal considerable inter-speaker differences in the effects of these factors on variation of acoustic-prosodic measures.

This paper is organized as follows. We first review evidence from prior studies for prosodic effects related to structural and referent-based factors in a variety of languages, motivating our focus on these factors in the analysis of Russian. Next, we provide an overview of the discourse and production data analyzed in the present study and formulate our hypotheses. Next, we report our experimental findings and discuss what these findings mean for our understanding of prosodic expression in read discourse in Russian. We then discuss our findings and conclude.

Prior work on structural, grammatical and semantic effects on prosody

Word order effects on prosody. Even more so than in fixed word order languages, in a 'free word order' language like Russian or Greek, the discourse status of a word, related to its semantic focus or information status, may be (probabilistically) reflected in its location in a dedicated clausal position (Clark and Clark, 1978; Baltazani 2003; Féry and Krifka 2008; Yokoyama 1986). Favored positions place discourse-given information before discourse-new information, which may be understood to facilitate sentence processing and ease comprehension (Clark and Haviland 1977; McDonald, Bock, Kelly 1993; Clifton and Frazier 2004). Thus, just as with prosodic encoding,

word order variability, when available, may be used to encode the information status of a word in relation to the discourse context.

Word order and prosody in Russian. In Hindi, Finnish, Romani, and, possibly other languages that exhibit discourse-motivated word order variability, a deviation from the canonical word order alters the prosodic properties of an utterance or phrase. To illustrate, Arvaniti and Adamou (2011) report that in the variety of Romani spoken in Komotini, narrow focus on a nominal expression may be realized by positioning the focused noun ex-situ, preverbally, and accenting it, along with deaccenting of the following verb. Vainio and Järvikivi (2006, 2007) report that in Finnish, a change in word order alters the intensity profile and tonal shape of an utterance and affect the perception of the ex-situ constituent as prominent. Similarly, Patil et al. (2008) report that in Hindi, ex-situ pre-verbal constituents have a greater pitch excursion and greater duration. Russian is known as a highly free word order language.[4] The ordering of words in a sentence serves a pragmatic function (Kallestinova 2007; Slioussar 2010, 2011b): when presented out of context, all word order permutations express the same propositional content, but once presented in discourse, they differ in pragmatic meaning (see Chapter 1 for an illustration). Previous work shows that in Russian examples such as this, a change in word order typically triggers a change in the acoustic-prosodic realization of the ex-situ word (Botinis, Themistocleous, Kostopoulos, Nikolaenkova 2005; Luchkina and Cole 2014). A growing body of research shows that in this respect Russian is similar to other languages, in which structural and prosodic means are engaged in parallel when

[4] According to corpus data from written Russian (Bivon 1971), 79% of 3-member sentences have the SVO order. The OVS order is the most frequent non-canonical word order in Russian and accounts for 11% of such sentences. SOV, VSO, VOS orders are possible and account for 1-4% of sentences each. Slioussar (2011) cites Sirotinina (1965) who estimated that 16-30% of sentences in written texts in Russian contain postverbal subjects. Lobanova (2011) reports that in a corpus of 300 sentences, 88% were SVO, 6% - OVS, 4% - OSV, 1.5% SOV and 0.5% - VOS. Derivation of non-canonical word orders in Russian has received considerable attention in the syntax literature (see Bailyn 1995, 2011, Babyonyshev 1996, Slioussar 2011, among others) and is outside the scope of this paper.

encoding referent status or greater information prominence in discourse (Greek: Baltazani 2003; Finnish: Vainio and Järvikivi 2006, 2007; Romani: Arvaniti and Adamou 2011; Hindi: Patil et al. 2008; Samoan: Calhoun 2015).

The interaction between argument linearization and prosodic expression in Russian (in comparison with Greek) was investigated by Botinis et al. (2005), who performed a comparative analysis of tonal and syntactic correlates of focus using elicited production data from over a hundred Russian speakers. Botinis and colleagues report that narrow focus productions in Russian involve a marked word order, OVS, and are clearly reflected in the 'tonal' or prosodic realization of an utterance, through a local tonal f0 range expansion in the vicinity of the focused constituent and the concurrent tonal compression in the vicinity of the unfocused material. Botinis et al.'s findings suggest that the tonal (f0) correlates of narrow focus in Russian are independent of the syntactic correlates, and that when used concurrently, prosodic and syntactic effects of focus may reciprocally reinforce each other. While Botinis et al.'s work provides a first important step to understanding the relationship between structural and prosodic cues in connected speech in Russian, it does not tease apart prosodic variation conditioned by constituent reordering, referent properties, and the pragmatic phenomenon of narrow focus.

Referent-based effects on prosody

Animacy and subjecthood. Apart from the influence that argument linearization may have on prosodic expression, the inherent properties of discourse referents, such as animacy, concreteness and definiteness, present other, potentially indirect, sources of prosodic variation. Across languages, an inherent referent feature animacy, through the notion of conceptual accessibility, is closely related to the thematic roles of the agent and patient for simple transitive verbs, as in the Russian example (2.1) (Bock and Warren 1985, Frazier and Clifton 1996; Bornkessel-Schlesewsky

and Schlesewsky 2009). Referent animacy is also related to the grammatical function of a referring expression, particularly, subjecthood, via the association of prototypical animate agents with grammatical subjects, and prototypical inanimate patients with grammatical objects (Comrie 1989; Kuperberg, Kreher, Sitnikova, Caplan, Holcomb 2007; Hoeks, Stowe, Doedens 2004; Kuperberg, Sitnikova, Caplan, Holcomb 2003; de Swart 2007). Cross-linguistic studies, e.g., Siewierska (1993), and Bresnan, Dingare, Manning (2001), have demonstrated that animate entities not only have a tendency to occur in higher syntactic positions, such as that of the sentence subject, but also tend to occur early in a sentence or clause. We therefore recognize that the animacy feature plays a role in determining the thematic role and grammatical function of a discourse referent in relation to a predicate, and that these factors may in turn influence word order.

Cross-linguistically, morphological (case marking) and syntactic (linearization) devices bear an effect on distinctions such as those between animate and inanimate referents, as well as between grammatical subjects and objects (Bornkessel-Schlesewsky and Schlesewsky 2009). Convergence between syntactic and conceptual representations, achieved when argument roles in the sentence are filled prototypically (i.e., when the subject referent is animate and the object referent is inanimate) has been found to facilitate sentence processing (Traxler, Morris, Seely 2002; Bornkessel-Schlesewsky and Schlesewsky 2009). In English, where word order flexibility is limited, animacy plays an important role in grammatical role assignment, but has no direct effect on linearization of major sentence constituents (Bock and Warren 1985; McDonald et al. 1993, among others). However, in a free word order language, animacy may affect the linear order of sentence constituents and determine argument linearization preferences, as shown in experimental work by Prat-Sala and Branigan (2000) and Verhoeven (2009, 2014). So called 'animate-first' effects, which translate into a preference for an animate referent to appear early in a sentence or

phrase, have been found in a number of languages with relatively flexible constituent order, including Greek and Turkish (Branigan and Feleiki 1999; Verhoeven 2014), Japanese (Tanaka, Branigan, Pickering 2005; Branigan, Pickering, Tanaka 2008), Spanish (Prat-Sala 1997), and German (van Nice and Dietrich 2003; Verhoeven 2014). To illustrate, in a sentence recall task conducted by Branigan and Feleki (1999), Greek speakers preferred the order of constituents in which the animate entity (subject or object) preceded the inanimate entity, regardless of the resulting word order (canonical SVO or non-canonical OVS). In a similar vein, in a sentence recall task conducted by Tanaka et al. (2005), Japanese speakers altered constituent order from OSV to SOV when the sentence subject was animate. In line with these findings, Lobanova (2011) examined the role of animacy on argument ordering in two Russian written corpora and found that animate nouns bear a strong association with the sentence-initial position, regardless of their grammatical function, as the subject in a canonical SVO sentence or the object in an object-initial OVS sentence.

Less is known about the potential for direct effects of animacy on prosodic variation. Antão et al. (2013) explored prosodic variation in relation to animacy and subjecthood in Brazilian Portuguese. Using a corpus of elicited speech, Antão et al. (2013) found that animacy and subjecthood significantly affected a number of prosodic parameters, including mean f0, f0 range, and duration, which were greater for animate referents and grammatical subjects. These findings are largely consistent with the view that animate referents and grammatical subjects have greater discourse salience than other types of referring expressions (McDonald et al. 1993; Mak, Vonk, Schriefers 2002; Traxler, Williams, Blozis, Morris 2005).

Relative accessibility and information status

Apart from the effects of animacy and subjecthood, prosodic variation may also arise due to the

information status of words and the relative accessibility of their referents for speech comprehension. Information structure relates the referent of a word to the preceding discourse, distinguishing the referent as given (theme), via prior mentions in discourse, or novel (rheme), introduced to discourse for the first time. These categories of information may occur together in a single utterance or phrase, partitioning the utterance into distinct information components, such as given information and novel information. While the given-novel (topic-comment, background-focus) dichotomy is central to many information structure approaches (e.g., Halliday 1967), more fine-grained distinctions between information categories have been proposed, including an influential tripartite distinction between new – inferable – evoked (given) information proposed by Prince (1981). More recently, Baumann and Riester (2012, 2013) proposed that information status effects can be considered in two forms: referential and lexical. *Referential information status* relates to the discourse status of the referent of a word, which is dynamic across the discourse, reflecting changes in the accessibility of the referent based on the prior discourse context. In cognitive accounts (Chafe 1976, 1994; Lambrecht 1994), the accessibility or givenness of discourse entities is described in terms of the activation costs associated with bringing these entities into the focus of the speaker's/hearer's attention. Referent accessibility may be viewed as a continuous measure, signaled by means of special morphological markers, word order, and/or prosody (Morgan et al. 1987). *Lexical information status* relates to the prior mentions of a lexical item, and along with lexical frequency (token frequency of a lexeme in the language), is known to have an effect on the ease of lexical access. The breakdown of information status into the referential and lexical levels is of particular relevance to this study, since it allows for a close-up analysis of referent-specific information status effects on prosodic expression in discourse.

Rich empirical evidence has been accumulated suggesting effects of information status, and word

predictability on acoustic measures of prosody (see, among others, Aylett and Turk 2004; Breen et al. 2010; Cole et al. 2011; Watson 2010; Baumann and Riester 2012, 2013; Cruttenden 2006; de Ruiter 2015). From research on many languages, including especially English, it is well known that the prosodic expression of a word reflects its information status (new or given) and its focus status (broad, narrow, contrastive). In English, the absence of morphological focus or topic markers and rigid constraints on word order mean that prosody is the primary mechanism for encoding this kind of discourse meaning. It is of interest then that prosodic encoding of discourse meaning is also identified for languages with 'free word order'. For example, f0 and duration are among the acoustic-prosodic correlates of perceived prominence in European Portuguese, Greek, Finnish, and Romanian—languages with some degree of variable word order (Frota 2002; Baltazani 2003; Vainio and Järvikivi 2006; Swerts 2007).

Referent-based effects on prosody in Russian

Russian has been shown to exhibit prosodic effects of referent information status in patterns of pitch-accenting, with accenting of novel information and deaccenting of given information (Neeleman and Titov 2009; Jasinskaja 2013). In addition, information status is reliably associated with the clausal positioning of a word (King 1995; Brun 2001). Specifically, the default (pragmatically neutral) pattern is for discourse-new information to occur clause-finally, while contrastive information may surface in various positions or remain *in-situ*.

Research questions

We have seen that the acoustic expression of prosody is variable and reflects structural factors such as word order, as well referent-based factors, related to semantic and grammatical properties, and information status. We have also seen that these referent-based factors interact with word order in languages such as Russian, which further complicates prosodic analysis. To date there is no

study that considers word order against the range of referent-based factors discussed above, which means that no study has fully addressed the interaction among these factors in the prosodic expression of a word. Without considering structural and referent-based factors together, it is not possible to answer even basic questions about the prosodic encoding of discourse meaning to determine if, for example, word order or acoustic prosodic cues are the primary means for encoding discourse meaning, if the two function in tandem, and to what extent factors related to the discourse referent of a word mediate in relationship between word order, prosody, and discourse meaning.

The goal of the present study is to test the effects of word order, animacy and grammatical function, and information status on the acoustic expression of prosody in Russian discourse. Our first research question is whether word order affects acoustic prosodic measures, such that words that are ex-situ relative to their canonical positioning (SVO) are acoustically distinct from canonically positioned words.

We address this question by examining acoustic-prosodic variation associated with two types of ex-situ positions, fronted (sentence-initial) and post-posed (sentence-final), while controlling for other potential sources of prosodic variation stemming from grammatical, semantic, and pragmatic properties of discourse referents. In line with the finding that ex-situ words in other free word order languages tend to be prosodically distinct (e.g., Vainio and Järvikivi 2006; Patil et al. 2008), we recognize that prosodic effects associated with an ex-situ position may be orthogonal to discourse meaning, and present the acoustic-prosodic 'aftermath' due to an ex-situ position in a sentence or phrase. Alternatively, such effects may be aligned with discourse meaning, as predicted by the evidence that the prosodic expression of an ex-situ word may cue its referent features and information status. Consistent with these outcomes we test the following hypothesis:

Hypothesis 1: *Deviations from canonical word order trigger prosodic effects that are word order-specific, i.e., structural in nature, and that are independent from the semantic and grammatical properties of a word or its information status.*

The second research question addressed in our study concerns the effect of referent-based factors and grammatical function of referring expressions on acoustic-prosodic measures. Prior work has established that these aspects of discourse meaning influence the linear ordering of sentence constituents and may receive special prosodic and structural expression in spoken and read discourse. We therefore ask how semantic, grammatical, and information-structural properties of discourse referents affect acoustic-prosodic measures in Russian, and specifically, if observable differences in prosodic expression can be traced to the animacy and subjecthood and information status of discourse referents. To answer this question, we evaluate, separately, acoustic-prosodic features associated with (1) animate vs. inanimate referents in conjunction with their grammatical function and (2) referent information status. We also test the interaction of word order and referent-based factors on prosodic variation. Consistent with the view that animate entities, grammatical subjects, and discourse-novel information are inherently more salient in discourse (e.g., Branigan et al. 2008; Bornkessel-Schlesewsky and Schlesewsky 2009; Breen et al. 2010; Baumann and Riester 2012), we test the following hypothesis:

Hypothesis 2: *The animacy, subjecthood, and referent information status of a word affect its acoustic-prosodic properties as independent factors and in their interaction with word order.*

The research questions addressed in this study will help uncover sources of prosodic variation in discourse in a free word order language like Russian. More detailed predictions about how these structural and referent-based factors affect acoustic prosodic measures are presented in the next section, where the coding schemes and measurements are introduced.

Analysis of the reading performance of fifteen Russian speakers will enhance our understanding of the conditions under which a speaker utilizes acoustic-prosodic resources in the expression of discourse meaning. It will additionally allow us to gauge, although tentatively, speaker-specific variability in two types of prosodic variation: structural, pertaining to word order, and referent-based.

Materials and method

The speech materials used in this study come from two published narratives, an excerpt from a biography and a complete folk tale (344 content words, 69 function words total). Two stylistically different texts were chosen to reflect more standard (text 1) and more colloquial (text 2) language use. Nouns occurring in the Nominative case were labeled as subjects and nouns occurring in the Accusative, Dative, and Instrumental cases were labeled as objects. The word orders encountered in text 1 include 29 SVO clauses, 3 OVS and 1 SOV clauses. The word orders encountered in text 2 include 25 SVO, 3 OSV, 2 VSO, 2 SOV, 4 OVS, 2 OV and 3 VS clauses. Such uneven distribution of word orders in the study materials is expected, given that SVO and OVS are the two most common word orders in Russian, and that SVO is the pragmatically neutral order compatible with all information structural configurations. The average sentence length in the corpus is 5.2 content words (SD=1.77); 18% of all content words (61 words, 45 referring expressions) occur in ex-situ positions associated with non-SVO orders. A nominal expression in the corpus was labeled ‘ex-situ’ based on the linear order of sentence constituents relative to the (main) verb of the sentence. Thus, all post-verbal subjects and pre-verbal objects were classified as ‘ex-situ’. Additionally, three instances of indirect Dative objects occurring after a direct object were also classified as ‘ex-situ’.

The information status of each referring expression in the corpus was annotated by one of the

authors (TL) and another native Russian speaker for a total of 259 content words. Inter-rater agreement (linearly weighted Kappa) between the annotators across texts was satisfactory: $\varkappa$=0.86, SE=0.03, α=0.05. Information status was classified based on Baumann and Riester's (2012, 2013) RefLex annotation scheme, rooted in Chafe's cognitive approach to information structure (Chafe 1976, 1994). As discussed in 2.3.2 above, RefLex allows for specification of information status at two qualitatively distinct levels, referential, pertaining to properties of the discourse referent of the word, and lexical, pertaining to the lexical choices a speaker makes to identify a discourse referent. Our interest in this study is on prosodic effects of word order and referent-based factors, so accordingly, we focus on referential level information status. Following Baumann and Riester (2012, 2013), we distinguish between four distinct classes of discourse referents: r-given, r-bridging, r-new, and r-unused[5], defined in Table 2.1. Each referring expression in the corpus was assigned to one of these classes. Each content word in the corpus was also annotated for its position in the sentence, being either in-situ or ex-situ relative to the canonical SVO order. More specifically, we distinguished between fronted sentence-initial and post-posed sentence-final ex-situ positions. All constituents in SVO sentences were coded as *in-situ*. Following Slioussar (2011b) and Ionin and Luchkina (under review), we treat both subject and object as ex-situ in the non-canonical OVS order; we therefore coded subject noun as *ex-situ post-posed* and object noun – as *ex-situ fronted* in the OVS sentences (cf. Bailyn 2003, 2004). Objects were coded as ex-situ fronted in the SOV and OSV non-canonical orders. Verbs and non-referring expressions in ex-situ positions were not included in subsequent analyses. The coding scheme for ex-situ constituents is illustrated in (2.2) and (2.3) taken from text 2.

[5] The category of r-unused information was only represented by 2 toponyms in the corpus. Because of being underrepresented, data from both r-unused words were excluded from statistical analyses.

Table 2.1. *Definitions of referential information categories (following Bauman and Riester 2012, 2013).*

	r-given	**r-bridging**	**r-new**	**r-unused**
Information status of discourse referents	anaphor coreferring with an antecedent in previous discourse (given status)	non-coreferring anaphor dependent on preceding context (inferable status)	referent/concept introduced to discourse for the first time (new status)	discourse-new item which is generally known (e.g, a toponym) (new status)

(2.2) V eto vremya po doroge shol kozyol. ex-situ post-posed

at this time along road walked-M goat-NOM

'At that time, a goat walked along the road.'

(2.3) ex-situ fronted

vody	v	kolodce	bylo	ne	mnogo
water-GEN	in	well	was	not	much

'There was not much water in the well.'

The animacy and grammatical function of discourse referents were coded jointly, by associating each nominal expression in the corpus with one of the four levels of the variable *AGRC* (AnimacyGRammaticalClass), which combines the animacy and grammatical function into a single category: animate subject; inanimate subject[6]; animate object; and inanimate object.

Acoustic features pre-processing and statistical modeling of production data

All materials were read aloud by 15 native speakers of Russian (9 females), ages 21-52. All speakers read the narratives silently and then were instructed to read them aloud in a lively naturalistic manner, as if addressing an audience. Recordings were made in a sound-proof recording booth using a Marantz PDM 750 solid state recorder and a head-mounted microphone.

[6] Only one grammatical subject in the corpus was inanimate. The category *inanimate subject* was therefore excluded from statistical analyses.

Recorded data were digitized at the sampling rate of 44.1 kHz and submitted to acoustic analyses. The acoustic measures of f0 range[7], mean intensity, and raw duration were taken from every syllable of each word in the corpus. All measurements were extracted automatically in Praat (Boersma and Weenink 2013). Fundamental frequency and intensity measures were taken from the center region of the vowel, excluding the 10-ms sub-regions from the vowel edges, in order minimize the influence of adjacent consonants at vowel onset and during inter-segmental transitions. Each f0 output was transformed to semitone values relative to a fixed value of 100 Hz. The acoustic measures of f0 range, mean intensity, and duration were then examined for their relationship to animacy, subjecthood, and the referent status of discourse entities, as well as word order.

For the purposes of the statistical analyses, intensity values were log transformed and duration values were subject to mean-centered coding using within-subject z-scores (Bush, Hess, and Wolford 1993). Normalization of acoustic-prosodic parameters was implemented to minimize variability between speakers which is due to individual characteristics and speech rate. A multivariate analysis of variance was fit to three dependent variables, *f0 range*, *mean intensity*, and *duration*. The following predictor variables entered the analysis: *AGRC* (levels: 'animate subject', 'animate object', 'inanimate object'), *IS REF* (referential information status, levels: 'r-given', 'r-bridging', 'r-new'), *word order* (levels: 'in-situ', 'post-posed', 'fronted'), and *speaker*. To control

[7] In the present study, we treat f0 range as a proxy of f0 excursion size, following Traunmüller and Eriksson (1995). f0 range was chosen as the f0 correlate following Botinis et al. (2005) who reported that compression and expansion of f0 range marked prominent (focused) constituents in their Russian and Greek data, regardless of word order manipulation. In prior research, f0 range was successfully tested as a correlate of focus and information status in various languages, including English (Xu and Xu 2005, Breen et al. 2010, among others), Hindi (Patil et al. 2008), Mandarin Chinese (Xu 1999). Other f0 measures available to us, including max, mean, and syllable-final f0, were not included into the list of predictors due to collinearity with f0 range and intensity or multiple instances of missing values.

for the effect of vowel aperture on intensity, *open vowel* (levels: 'open vowel' and 'other') was also added as a control factor to account for possibly greater intensity of unreduced open vowels in stressed syllables (e.g., word-final vowel in the word Moskvá). Predictor variables and their levels are summarized in Table 2.2.

Table 2.2. *Fixed effects used in multinomial regression model. Reference level for each factor is highlighted in bold.*

fixed effects	Levels
word order	**in-situ**
	ex-situ, fronted
	ex-situ, post-posed
information status	**r-given**
	r-bridging
	r-new
AGRC	object, animate
	object, inanimate
	subject, animate
vowel height	**stressed open**
	other

The omnibus MANOVA returned tests of significance for the overall model using four different multivariate criteria, Wilks' lambda, Lawley-Hotelling trace, Pillai's trace, and Roy's largest root. The model was overall highly significant, regardless of the type of the multivariate criterion (all p values <.001, e.g., using Wilks lambda, $F(23, 8171)=475.1$, $p<.001$). The MANOVA revealed that the control variable *open vowel*, although significant overall, accounted for the least amount of variance in the dependent variables ($F(1, 8169)=21.4$, $p<.001$, $\eta^2<.01$). Predictor variable *speaker* accounted for the largest amount of variance in the dependent variables and the largest effect size in the model ($F(14, 8171)=759.2$, $p<.001$, $\eta^2=.53$). Predictor variables *word order* ($F(3, 8170)=55.64$, $p<.001$, $\eta^2=.1$), *IS REF* ($F(3, 8171)=52.05$, $p<.001$, $\eta^2=.1$), and *AGRC* ($F(3, 8171)=73.81$, $p<.001$, $\eta^2=.1$) in turn, proved to be robust, however, with small effect size.

Following the multivariate analysis of variance, a multivariate linear regression model[8] was fit to the data. Consistent with MANOVA, *mean intensity, duration,* and *f0 range* were introduced as dependent variables. *AGRC, word order, IS REF* and *vowel height* were included as fixed effects. *Speaker* was introduced as a random effect. The model also included interactions between *AGRC, word order, and IS REF*. The proportion of variance in each dependent variable that is explained jointly by the independent variables and the interaction terms was estimated separately for each acoustic-prosodic parameter in the model as follows: mean intensity: F=355.98 p<.001, R^2(adjusted)=47.82; duration: F=399.67 p<.001, R^2(adjusted)=50.7; f0 range: F=18.11, p<.001, R^2(adjusted)=.45.

Predicted effects of structural and referent-based factors on acoustic-prosodic measures

Having introduced the acoustic prosodic measures and the coding scheme for the predictor variables, we are now able to formulate the specific effects that are predicted by our two hypotheses, as stated earlier. For all of the predictions below, we consider greater values for f0 range, mean intensity, or duration as an *augmented* expression of a word's prosodic features, and conversely, diminished values of the same measures are considered as a *reduced* expression of a word's prosodic features. In this manner, we examine the effects of structural and referent-based factors in augmenting or reducing the acoustic-prosodic expression of a word.

(4) Predicted effects of structural and referent-based factors on acoustic-prosodic measures*:*

[8] In a multivariate regression analysis, several dependent variables are jointly regressed on the same independent variables. The individual coefficients and standard errors produced by a multivariate regression are identical to those that would be produced by sequential linear regressions estimating each regression equation separately. The difference is that being a joint estimator, a multivariate regression also estimates the between-equation covariances, allowing for meaningful direct comparison of coefficients across equations.

(a) [Hypothesis 1: structural effects on prosody] In line with findings from prior work on prosodic effects of non-canonical word order, we predict prosodic augmentation in Russian for words in ex-situ positions relative to words located in-situ.

(b) [Hypothesis 2: effects of animacy and subjecthood on word order] Here we consider predictions for effects of animacy and subjecthood on word order. We are interested in factors affecting word order because, due to the hypothesized effects of word order on acoustic-prosodic measures (see prediction a), any effect on word order may indirectly influence prosody. Following Lobanova's (2011) finding that animate nouns in Russian tend to occur sentence-initially, whereas inanimate nouns tend to occur sentence-finally, independent of their grammatical function, we predict a greater rate of fronting for animate objects and a greater rate of post-posing for inanimate subjects. Our corpus provides very few instances of inanimate subjects, so the second prediction here cannot be evaluated in our materials.

(c) [Hypothesis 2: effects of animacy and subjecthood on prosody] Because of greater inherent salience attributed to animacy and subjecthood, we predict prosodic augmentation for words with animate referents and for grammatical subjects, compared to words with inanimate referents and grammatical objects.

(d) [Hypothesis 2: effects of referent information status on word order] Here we consider predictions for effects of referent information status on word order, which again, is of interest due to the effect of word order on prosodic expression. Following findings from prior research showing a cross-linguistic preference for given information to precede novel information in a sentence or phrase, and considering the word order variability in Russian, we predict that words will be assigned to sentence positions depending on their information

status. More specifically, and consistent with the preference for discourse-given information to occur sentence-initially and discourse-new information to occur sentence-finally, we predict that words labeled as r-given will be fronted more often than post-posed and that words labeled as r-new will be post-posed more often than fronted.

(e) [Hypothesis 2: effects of referent information status on prosody] Consistent with the cross-linguistic preference for discourse-new information to be pitch-accented, and for discourse-given information to be deaccented, we predict prosodic augmentation for words whose referents are r-new compared to words with other information status labels. Further, given the relatively higher accessibility and lower information value associated with r-given and r-bridging information, we predict reduced prosodic expression for words whose referents are r-given and r-bridging.

Results

Table 2.3. lists beta coefficients for significant main effects and interactions obtained in the multivariate regression analysis organized by acoustic-prosodic dependent variable. Beta coefficients listed in Table 2.3. allow for direct comparisons of effects across regression equations. In what follows, we discuss these results based on the hypotheses and predictions formulated above. We first report the effect of word order on prosodic variation in the read productions of the corpus. Then, per (4b) and (4c), we examine prosodic variation in conjunction with two related properties of discourse referents, animacy and subjecthood. Next, we present an interim summary of these findings before turning to the results for information status. Finally, we examine prosodic variation in conjunction with information status of discourse referents, per (4d) and (4e). For each of these analyses, we present significant main effects and interactions based on the read production data from all fifteen speakers.

Table 2.3. *Beta coefficients for significant main effects and interactions in the multinomial regression model. Shaded cells indicate lack of significant effect.*

		β coefficients for acoustic-prosodic parameters		
fixed effect & interactions	levels of fixed effect	f0 range	mean intensity	duration
AGRC	objectA	-0.61	-3.24	
	objectI	-0.34	-1.67	-0.13
WO	ex-situ fronted	1.10	3.13	
	ex-situ post-posed	-0.68	-0.02	0.60
IS REF	New	-0.26	2.03	-0.34
	bridging	-0.27	0.04	-0.16
*AGRC*wo*	ex-situ fronted*objectA			0.49
	ex-situ post-posed*subjectA		0.59	0.72
*IS REF*wo*	ex-situ fronted*r-bridging	-1.71	-0.76	-0.31
	ex-situ post-posed*r-bridging	-0.31	-0.40	-0.45
*IS REF*wo*AGRC*	r-new*ex-situ post-posed*subjectA	1.58	5.73	0.84
	r-new*ex-situ fronted*objectA		5.62	0.43
	r-new*ex-situ fronted*objectI		3.76	0.99

all effects significant at $p<.05$

Recall that in the omnibus MANOVA model, the factor *speaker* accounted for more variation in the prosodic parameters examined in this study than any other predictor did. In the post-hoc analyses of the interactions between *word order, AGRC,* and *IS REF*, we therefore analyze production of individual speakers and illustrate inter-speaker differences in prosodic variation in response to these effects. We report post-hoc analyses of individual speaker productions with focus on *word order* and *AGRC,* as well as post-hoc analyses with focus on word order and referent information status.

Ex-situ position. Our first hypothesis concerns changes in prosodic expression related to a discourse-driven change in word order in Russian. Table 2.4. presents summary statistics for the acoustic-prosodic measures of interest computed for production data from all speakers across three sentence position types, in-situ, ex-situ fronted, and ex-situ post-posed.

Table 2.4. *Summary statistics for prosodic parameters duration (ms), mean intensity (dB) and f0 range (ST) across levels of Word Order based on the read production data from fifteen speakers.*

duration (ms)				
	n words	mean	SD	range
Fronted	28	75.88	38.33	264.34
in-situ	208	74.57	39.57	362.12
post-posed	33	82.42	41.83	233.65
mean intensity (dB)				
	n words	mean	SD	range
Fronted	28	77.81	5.75	32.85
in-situ	208	76.50	6.34	58.68
post-posed	33	75.60	6.69	42.32
f0 range (ST)				
	n words	mean	SD	range
Fronted	28	4.69	2.47	13.64
in-situ	208	2.33	2.47	25.85
post-posed	33	2.36	2.67	25.05

Descriptively, the co-variate *Word Order* has an effect on acoustic-prosodic expression, as seen in Table 2.4. Segment duration in post-posed words demonstrated, on average, an increase of 10 ms; similarly, the f0 range in ex-situ fronted words showed an increase of about 2 ST.

In the multivariate regression analysis, acoustic-prosodic features of ex-situ fronted and post-posed words were estimated relative to the prosodic characteristics of in-situ words. Controlling for the independent variables *AGRC* and *information status (IS REF)*, the dependent variables *mean intensity* and *f0 range* were greater for words fronted relative to their canonical position (*mean intensity*: t=5.74, p<.001; f0 range: t=3.70, p<.001). On the other hand, smaller *f0 range* (t=-2.69, p<.01) but greater *duration* (t=3.31, p=.001) were associated with words post-posed relative to their canonical position. These results partially support prediction (4a).

Animacy and subjecthood. Taking into account the interrelatedness of referent animacy and argument linearization reported in earlier work on Russian (Lobanova 2011) and other languages

with variable word order (Branigan and Feleki 1999; Tanaka et al. 2005), we first evaluated the rate of fronting and post-posing of a word in relation to the animacy of its referent. The observed distribution of *AGRC* categories across in-situ, fronted and post-posed sentence positions is illustrated in Figure 2.1.

Figure 2.1. *Distribution of AGRC across sentence positions. Y-axis: levels of AGRC; X-axis: percent occurrences.*

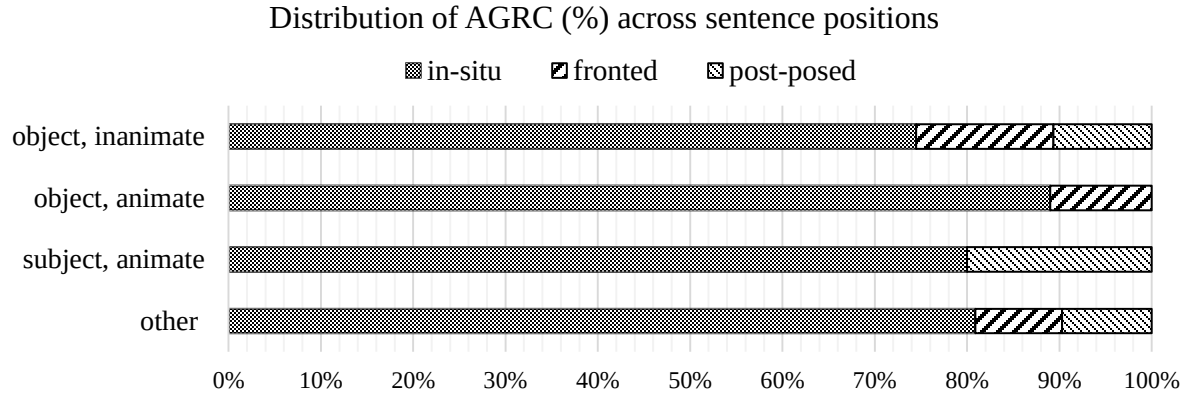


Figure 2.1. demonstrates that approx. 20% of animate subjects in the corpus occur sentence-finally. Due to the insufficient number of inanimate subjects in the corpus, prediction (4b) cannot be tested. We therefore proceed to evaluate the effect of animacy based on the ex-situ occurrence of objects. Figure 2.1. also demonstrates that approx. 12% of animate objects and 14% of inanimate objects are fronted relative to their canonical position. We observed highly comparable rates of fronting for animate and inanimate objects in the corpus, hence prediction (4b) was not borne out—we do not find differences in word order due to a word's animacy.[9]

[9] An anonymous reviewer rightly points out that *AGRC* holds a meaningful relationship with referent information status. The subject-object asymmetry with regards to animacy and given vs. novel information status of the referent is evident from the fact that animate subjects in the corpus are more often discourse-given than discourse-new, whereas inanimate objects are more often discourse-new than discourse-given (Pearson χ^2=263.55, p<.001). One reason underlying this asymmetrical distribution of subjects and objects across information classes may be the overall greater conceptual accessibility of (animate) subjects (Bock and Warren 1985, Branigan et al. 2008, Bornkessel-Schlesewky and Schlesewsky 2009), which makes them better prototypical topics, due to greater overall saliency in discourse. Additionally, in many languages, grammatical subjects appear sentence-initially and by doing so maintain desired given before new information structural configuration.

Of particular interest for this work is prosodic variability specific to grammatical function and animacy of discourse referents. Hence, read productions of the corpus were examined for prosodic variation specific to the animacy and grammatical function of discourse referents.

Table 2.5. *Summary statistics for prosodic parameters duration (ms), mean intensity (dB) and f0 range (ST) across levels of AGRC based on the read production data from fifteen speakers.*

duration (ms)				
	n words	mean	SD	range
object, animate	112	79.58	44.64	253.45
object, inanimate	80	75.75	36.70	331.57
subject, animate	66	84.45	43.20	356.74
non-nominal expressions	85	73.34	38.88	341.49
mean intensity (dB)				
	n words	mean	SD	range
object, animate	14	75.36	6.34	41.51
object, inanimate	15	76.20	6.36	48.18
subject, animate	29	78.15	5.83	35.25
non-nominal expressions	85	76.46	6.37	58.68
f0 range (ST)				
	n words	mean	SD	range
object, animate	14	2.36	2.86	24.03
object, inanimate	15	2.27	2.42	22.12
subject, animate	29	4.80	2.78	25.03
non-nominal expressions	85	2.31	2.40	25.85

Table 2.5. presents summary statistics for the acoustic-prosodic measures f0 range, mean intensity, and duration computed for production data from all speakers across the levels of *AGRC*, 'subject animate', 'object animate', 'object inanimate'. Descriptively, the co-variate *AGRC* has an effect on acoustic-prosodic expression, as seen in Table 2.5. Segment duration in the corpus was greatest for animate subjects; duration increment was, on average, 5 ms for animate subjects compared to animate objects, and 11 ms for animate subjects compared to the non-nominal expressions in the

corpus. Animate subjects, additionally, had approx. 3 dB higher mean intensity than animate objects and approx. 2.5 ST greater f0 range than animate objects. Despite the null effect of animacy and grammatical function on word order, results of the regression analysis confirm that animacy and subjecthood do have a robust direct effect on acoustic-prosodic measures, confirming prediction (4c).

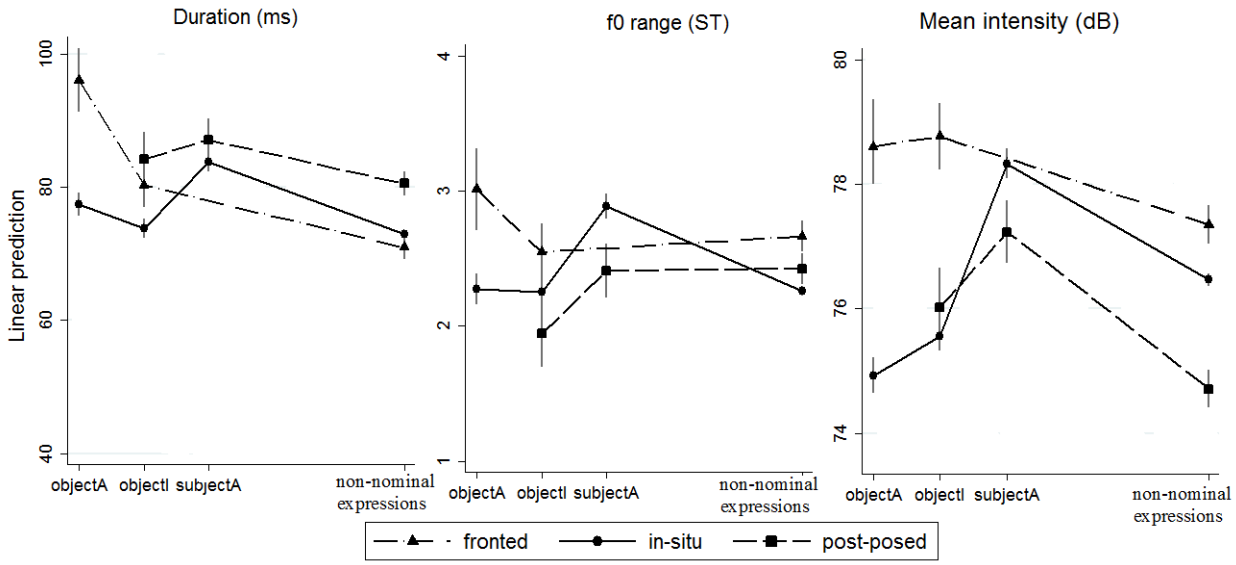
The animacy effect: Recall that animacy and grammatical function are crossed as levels of one variable, *AGRC,* in the regression model summarized in Table 2.5. This coding scheme allows for direct comparison of the acoustic-prosodic outcomes for all levels of *AGRC* relative to the baseline level of animate subjects, including the critical comparisons between animate and inanimate objects, and animate subjects and animate objects. We use these comparisons to evaluate, separately, effects of animacy and subjecthood on the acoustic-prosodic expression. Results reveal that the animacy of a word's referent significantly predicts variation in vowel duration; specifically, we found that duration was reduced for inanimate objects (t=-3.38, p<.01), but not for animate objects (t=-.8, p=.43).

The subjecthood effect: As shown in Table 2.5., regardless of the word order, *mean intensity* and *f0 range* were significantly reduced for every category of *AGRC* relative to that of animate subjects. Specifically, animate subjects had greater mean intensity than animate objects (t=-9.78, p<.001), and inanimate objects (t=-5.34, p<.001). Similarly, f0 range was significantly greater for animate subjects than objects, animate (t=-3.39, p<.005) or inanimate (t=-4.37, p<.001).

The mean values of the acoustic-prosodic variables for each level of *AGRC* are plotted in Figure 2.2. and reflect production data from all speakers in the sample. Significant two-way interactions between *AGRC* and *word order* illustrated in Figure 2.2. included greater mean intensity for ex-situ post-posed animate subjects (t=4.74, p<.001) and greater duration for fronted animate objects

(t=2.62, p<.01) and post-posed animate subjects (t=3.43, p=.001).

Figure 2.2. *Predicted marginal effects at the means with 95% CIs for parameters duration,, mean intensity, and f0 range across levels of AGRC and Word Order. Data from fifteen speakers. Y-axis: acoustic-prosodic parameter; X-axis: levels of AGRC.*



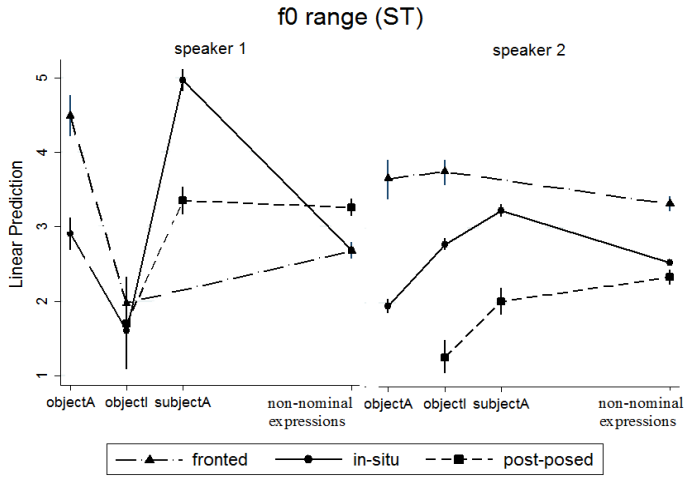
Inter-speaker variability and interaction with word order

Further exploration of the read productions of the narratives reveals that individual speakers differ in terms of which acoustic parameters they use to express prosodic features, which in our data gives rise to differences in the observed effects of structural and referent-based factors on acoustic-prosodic measures, and in their interaction. Individual speaker effects were tested by computing predicted marginal effects at the means[10] for the acoustic-prosodic parameters duration, f0 range, and mean intensity for *AGRC* and *word order,* using production data from each individual speaker summarized in Appendices 1 and 2. In what follows, for each acoustic-prosodic parameter of

[10] Marginal effects [MEs] post-estimations used in this study show how change in the variable of interest is related to change in one or more covariates. The MEs for categorical variables used in our post-estimation analyses show how a prosodic parameter changes as the categorical covariates *AGRC* and *Word Order* change from reference level to another level, after controlling for the other variables in the model (Cameron and Trivedi 2013).

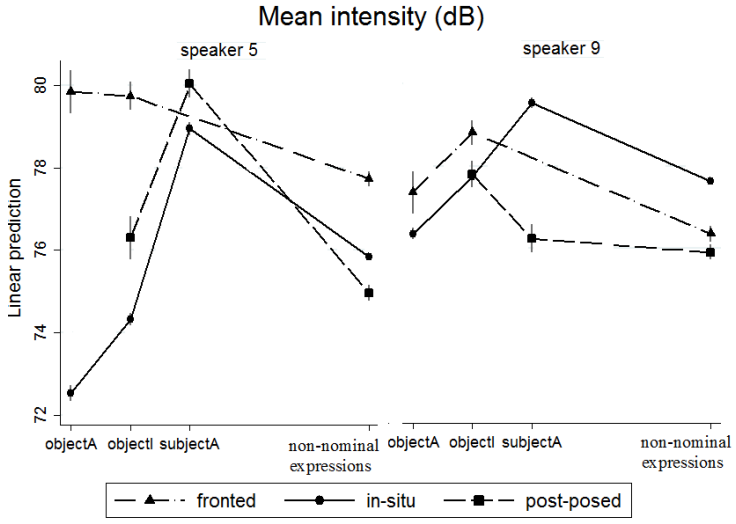
interest we select two speakers who differ in the effects of *AGRC* and *Word Order* on prosody. Specifically, for each acoustic-prosodic measure we plot the predicted marginal effects for two speakers, identified by their ID numbers, one who actively deploys that parameter, referred to as a high profile speaker, and one who deploys that parameter to a lesser degree, referred to as a low profile speaker.

Figure 2.3. *Predicted marginal effects at the means with 95% CIs for parameter f0 range across levels of AGRC and sentence position. Data from speakers 1 and 2. Y-axis: f0 range (ST); X-axis: levels of AGRC.*



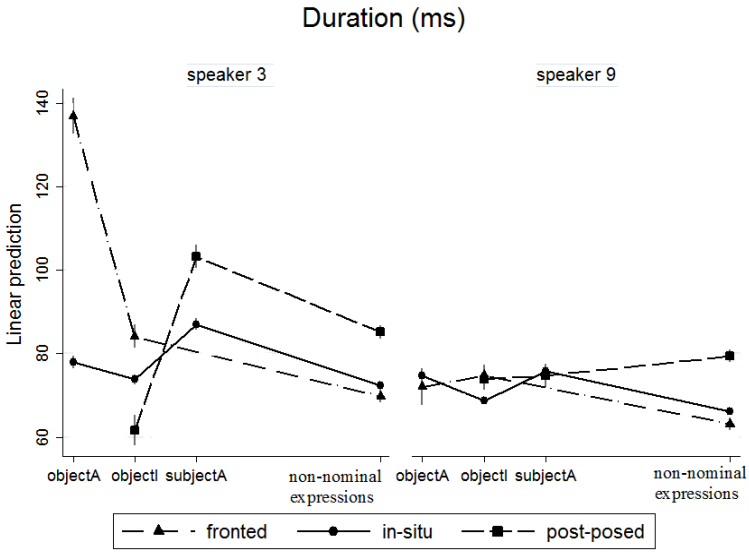
f0 range co-varied with referent animacy in the productions of eight speakers, increasing by 2-3 ST for animate referents relative to inanimate referents. Four speakers showed an increase in f0 range for animate objects that is 2-3.5 ST higher than for inanimate objects. Visual examination of individual speakers' plots of predicted marginal effects revealed that *Word Order* interacted with *AGRC* for seven speakers. The interaction pattern was uniform across these speakers: the animacy effect was greater for words occurring in ex-situ fronted positions but reduced for words occurring in-situ. Figure 2.3. demonstrates plots of predicted marginal effects for f0 range for the high profile speaker 1 (showing animacy, subjecthood, and word order effects) and the low profile speaker 2 (showing subjecthood and word order effects).

Figure 2.4. *Predicted marginal effects at the means with 95% CIs for parameter mean intensity across levels of AGRC and sentence position. Data from speakers 5 and 9. Y-axis: mean intenisty (dB); X-axis: levels of AGRC.*



Greater mean intensity in the corpus was observed in ex-situ fronted positions; this boost was most visible in productions of five speakers when the fronted constituent was an animate object.

Figure 2.5. *Predicted marginal effects at the means with 95% CIs for parameter duration across levels of AGRC and sentence position. Data from speakers 3 and 9. Y-axis: duration (ms); X-axis: levels of AGRC.*



Mean intensity also co-varied with subjecthood, with significant increase of 3-7 dB observed in productions from six speakers. Visual examination of individual speakers' plots of predicted marginal effects revealed that the animacy effect was consistently conservative and could be

cancelled out by the effect of word order: For three speakers there was no increase in intensity for post-posed animate nouns. Figure 2.4. demonstrates plots of predicted marginal effects for mean intensity for a high profile speaker 5 (subjecthood and word order effects) and a low profile speaker 9 (no significant effects of interest).

Duration reliably co-varied with referent animacy in the productions of ten speakers. On average, an increase of 10-70 ms was observed for words that are animate objects, compared to inanimate objects. Visual examination of individual speakers' plots of predicted marginal effects revealed that for eight speakers, the animacy effect was particularly apparent in the fronted, sentence-initial position. Duration also reliably co-varied with subjecthood for six speakers, although the increase in duration associated with grammatical subjects was considerably smaller, ranging between 5 and 30 ms. Figure 2.5. demonstrates plots of predicted marginal effects for duration for a high profile speaker 3 (animacy and subjecthood effects) and a low profile speaker 9 (no effects of interest).

Interim discussion

In the first series of analyses, we examined the effects of word order and two highly related properties of discourse referents, animacy and subjecthood, on acoustic-prosodic expression in the read production of two published narratives in Russian. We found that dislocating a word from its canonical position in a sentence or clause, not uncommon for Russian, triggers a chain of prosodic effects specific to the new surface position of that word. Analysis of the production data from fifteen native Russian speakers revealed that words fronted relative to their canonical position have greater mean intensity and f0 range, possibly reflecting articulatory strengthening at the left edge of a sentence or clause, a position typically aligned with the start of a prosodic-phrase. Post-posed sentence-final words, occurring at the likely end of a prosodic phrase, have smaller f0 range but greater duration, indicative of a final lengthening phenomenon.

These results confirm our first hypothesis stating that a deviation from the canonical word order in discourse triggers prosodic effects which we refer to as structural, i.e., not directly related to semantic and grammatical referent properties or referent information status.

Next, we examined effects of referent-based factors, looking first for effects of animacy and subjecthood on word order. If such effects are found, then we would understand word order to be a mediator for effects of animacy and subjecthood on acoustic-prosodic measures. We predicted that word order may act as a mechanism maintaining the optimal distribution of animate and inanimate referents across sentence positions, wherein an animate subject precedes an inanimate object. This prediction was not borne out: 20% of the animate subjects in the corpus occurred sentence finally (as in $OVS_{+animate}$ order) and slightly more inanimate objects (approx. 14%) occurred sentence-initially (as in $O_{-animate}VS$ order) than animate objects (approx. 12%). These results suggest that for the narratives used in this study, animacy is an unlikely factor driving the change in word order.

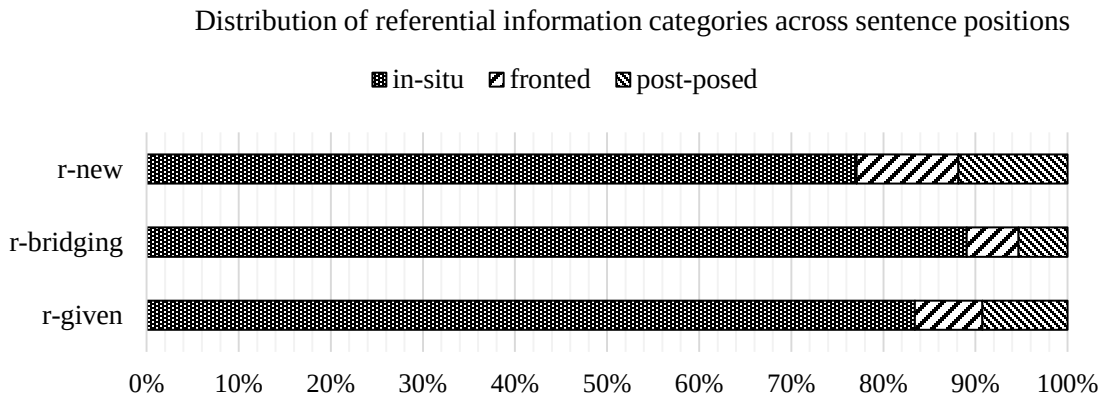
We further tested for direct effects of animacy and subjecthood on acoustic-prosodic measures. Analysis of the production data reveals greater f0 excursion and intensity for words that are subjects and for words with animate referents, and greater duration for words with animate referents. Follow-up analyses of individual speakers' production data lead us to qualify the effects of animacy and subjecthood on acoustic-prosodic expression: considerable inter-speaker variability in the use of acoustic-prosodic parameters in relation to animacy and subjecthood suggests that observed effects represent options rather than rule-like principles governing the reading performance of the study participants. Besides the finding of individual speaker differences, we uncovered multiple cases of interaction between animacy, subjecthood, and word order that further modulate acoustic-prosodic variation. In particular, an ex-situ fronted position is

where we observed the most dramatic increase in duration and f0 range for animate referents.

In summary, we have presented initial evidence for concurrent and incremental acoustic-prosodic effects of structural and referent-based features during the read production of discourse. In the following section, we address prosodic variation attributable to referent information status, reflecting changes in the readers' knowledge state as they progress through a text or narrative.

The information status of all content words in the corpus as entered in statistical analyses was classified using three referent information categories from the RefLex scheme (see Table 2.1.). The choice of the RefLex categories was based on how well-represented they were in the study materials. We begin by examining the distribution of the information categories r-given, r-new, and r-bridging, across in-situ, ex-situ fronted, and ex-situ post-posed sentence positions in the corpus shown in Figure 2.6. The first part of prediction (4d) was not borne out: words labeled as r-given were equally likely to be fronted (approx. 7% of all r-given words) or post-posed (approx. 9% of all r-given words), while occurring in-situ 83% of the time. The second part of prediction (4d) was partially borne out: Figure 2.6. shows that words labeled as r-new accounted for more non-canonical occurrences in the corpus than words representing any other information status. Discourse-new information accounted for the largest number of post-posed ex-situ positions (62.4% of all post-posed occurrences), as well as fronted ex-situ positions (62.9% of all fronted occurrences). The observed distribution, however, does not support the tendency for discourse-new information to be post-posed, as predicted by (4d).

Figure 2.6. *Distribution (%) of referential information categories across sentence positions. Y-axis: Information status. X-axis: % occurrences.*



We turn now to examine direct effects of referent information status on acoustic-prosodic variation, in the three information categories that are well-represented in our corpus: r-given, r-new, and r-bridging. Table 2.6. presents summary statistics for the acoustic-prosodic measures of interest computed using production data from all speakers across these three levels of information status.

Table 2.6. *Summary statistics for prosodic parameters duration (ms), mean intensity (dB) and f0 range (ST) across levels of IS REF based on the read production data from fifteen speakers.*

duration (ms)				
	n words	mean	SD	range
r-given	84	81.63	41.02	360.73
r-bridging	53	67.01	31.70	222.56
r-new	121	74.79	40.91	341.55
mean intensity (dB)				
	n words	mean	SD	range
r-given	84	76.49	6.43	50.28
r-bridging	53	77.23	5.88	40.71
r-new	121	76.31	6.43	58.68
f0 range (ST)				
	n words	mean	SD	range
r-given	84	2.62	2.68	24.04
r-bridging	53	2.10	2.16	24.80
r-new	121	2.31	2.47	25.85

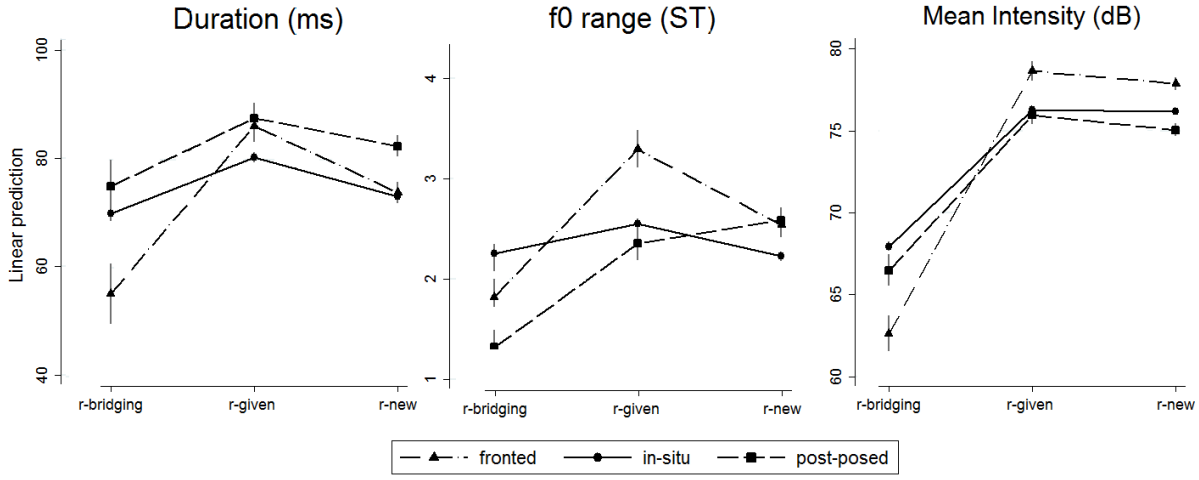
Descriptively, little variation is visible in the mean values of acoustic-prosodic predictors for each referent information status; however, we observe that r-given referents have, on average, greater duration. Additionally, the ranges of f0 excursion and intensity means are slightly greater for r-new words, as seen in Table 2.6.

Results of the multivariate regression analysis provide partial support for the predictions in (4e).

Discourse-new status: Relative to the reference category of r-given and controlling for word order and animacy/subjecthood, words labeled as r-new had significantly higher mean intensity (t=2.32, p<.05). However, no further evidence for prosodic augmentation of discourse-new information was obtained. In fact, f0 range and duration measures were systematically greater for r-given words in the corpus (t=-3.84, p=.001 for *f0 range* and t=-5.36, p<.001 for *duration*), regardless of word order.

Discourse-bridging status: Consistent with prediction (4e), relative to the reference category of r-given and controlling for word order and animacy/subjecthood, discourse-bridging words had reduced *f0 range* (t=-4.97, p<.001), and reduced *duration* (t=-8.77, p<.001).

Figure 2.7. *Predicted marginal effects at the means with 95% CIs for parameters duration,, mean intensity, and f0 range across levels of IS REF and Word Order. Data from fifteen speakers. Y-axis: acoustic-prosodic parameter; X-axis: levels of IS REF.*



The mean values of acoustic-prosodic measures for each level of referent information status is plotted in Figure 2.7., based on production data from all speakers in the sample. Analysis of two-way interactions between referent information status and word order illustrated in Figure 2.7. reveals that highly inferable r-bridging referents, when occurring ex-situ, underwent further reduction in *f0 range* (ex-situ fronted: t=-3.1, p<.005; ex-situ post-posed: t=-3.4, p<.001),*duration* (ex-situ fronted: t=-2.79, p<.01; ex-situ post-posed: t=-2.05, p<.05), and *mean intensity* (ex-situ post-posed: t=-5.98, p<.001).

As Table 2.3. shows, *IS REF* also entered into a number of significant 3-way interactions with *AGRC* and *word order*. Specifically, discourse-novel post-posed animate subjects were prosodically augmented in terms of each acoustic-prosodic parameter of interest (*f0 range*: t=3.57, p<.001, *mean intensity*: t=4.80, p<.001; *duration*: t=3.43, p=.001). Discourse-novel fronted objects were also prosodically marked, as follows: animate objects had greater *mean intensity* (t=2.45, p<.01) and *duration* (t=2.62, p<.01), and inanimate objects had greater *mean intensity* (t=3.54, p<.001).

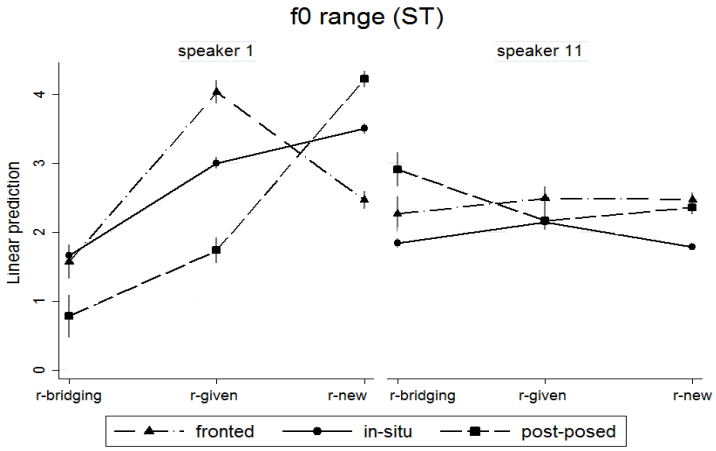
Next, we complement these results with post-estimation analyses of individual speakers' contribution to the variance in the acoustic-prosodic parameters of interest.

Inter-speaker variability and interaction with word order (continued)

We again notice individual speaker differences in the acoustic parameters that show variation due to referent-based factors, here considering referent information status. To examine individual differences in main effects of information status, and its interaction with word order, we computed predicted marginal effects at the means for the acoustic-prosodic parameters duration, f0 range, and mean intensity for *IS REF* and *Word Order,* using production data from each individual speaker also summarized in Appendix C. As above, for each acoustic-prosodic parameter of

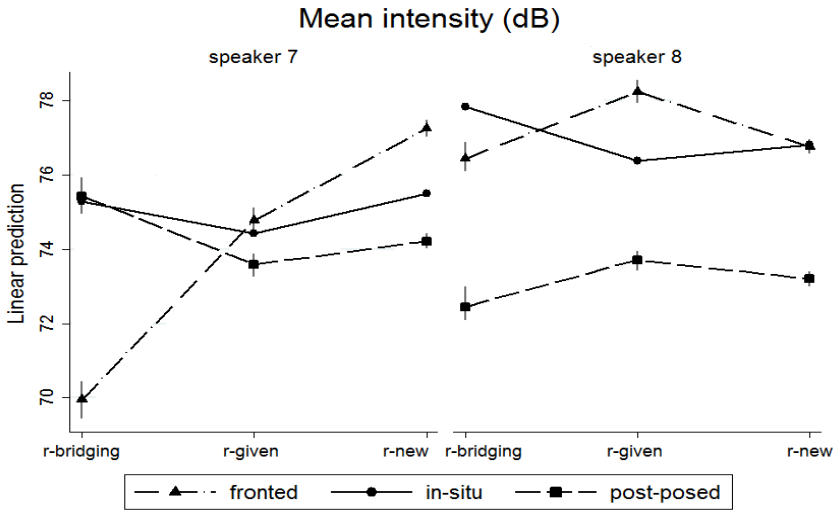
interest, we plot the predicted marginal effects for two speakers, identified with their ID numbers, one who actively deploys that parameter, referred to as a high profile speaker, and one who deploys that parameter to a lesser degree, referred to as a low profile speaker.

Figure 2.8. *Predicted marginal effects at the means with 95% CIs for parameter f0 range across levels of information status and sentence position. Data from speakers 1 and 11. Y-axis: f0 range (ST); X-axis: levels of information status.*



f0 range was greater for r-given information, contrary to what we predicted in (4e). Visual examination of the individual speakers' plots of predicted marginal means reveals that f0 range was dually affected by the information status of a word and its position in the sentence or phrase. Figure 2.8. shows that in the production of the high profile speaker 1, f0 range was largest for fronted r-given words and post-posed r-new words. While this pattern of results holds for ten speakers in our sample, for nine out of these ten speakers, the increase in the f0 range associated with post-posed discourse-new information was more modest than that observed for fronted discourse-given words, hence the direction of the main effect. The plot for a low profile speaker 11 (see Figure 2.8.) demonstrates a lack of an effect of information status or word order.

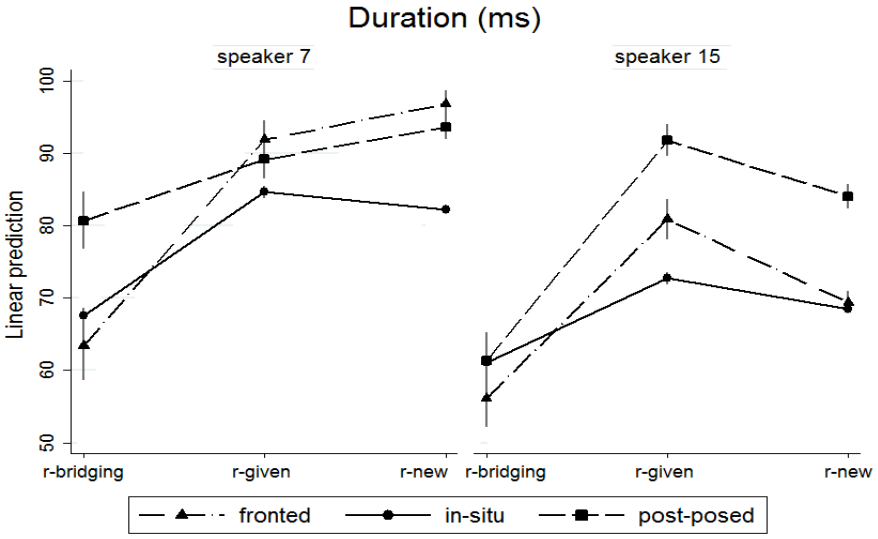
Figure 2.9. *Predicted marginal effects at the means with 95% CIs for parameter mean intensity across levels of information status and sentence position. Data from speakers 7 and 8. Y-axis: mean intensity (dB); X-axis: levels of information status.*



Mean intensity in the corpus served as the only robust cue for r-new information, with modest but significant increases of 5-7 dB observed in the productions of eleven speakers. Visual examination of the individual speakers' plots reveals that the effect was consistently conservative. Figure 2.9. demonstrates plots of predicted marginal effects for mean intensity for a high profile speaker 7 (main effect of *IS REF*) and a low profile speaker 8 (main effect of *Word Order*).

Duration was consistently greater for r-given information than for r-bridging or r-new information. The plot illustrating productions of speaker 15 shown in Figure 2.10. is highly representative of this effect and characterizes eleven speakers in our sample. The four remaining speakers did not demonstrate an increase in duration for r-given information. In three speakers' productions, fronted discourse-new information had significantly greater duration (see the plot illustrating production of speaker 7).

Figure 2.10. *Predicted marginal effects at the means with 95% CIs for parameter duration across levels of information status and sentence position. Data from speakers 7 and 15. Y-axis: duration (ms); X-axis: levels of information status.*



Discussion

As discourse unfolds, information which may be perceived as less accessible or novel at the beginning of a narrative undergoes an increase in accessibility and changes its status as the reader progresses through the narrative. In this study, we ask if the information status of a word's referent, which reflects the dynamic knowledge state of the reader, is reflected in the placement of the word in the sentence or phrase, and if information status has an effect on prosodic expression in read production of discourse. Using a corpus of two published narratives in Russian, a free word order language, we first examined the distribution of content words falling into distinct referent information classes across in- and ex-situ sentence positions. We used the information structure annotation scheme proposed by Baumann and Riester (2012, 2013), to reflect the dynamic knowledge state of the reader as they progress through discourse. Distinct referent information status categories we investigated are new, given, and bridging or inferable. We found that the distribution of these information types across in-situ and ex-situ positions in a sentence or phrase is indicative of their relative accessibility for the reader. Specifically, the relative accessibility of

a word, as indexed by a RefLex category, is negatively associated with an ex-situ position, as evident from the finding that highly inferable discourse-bridging words have the most in-situ occurrences in the corpus, whereas discourse-new words, introducing new discourse referents and considerably less accessible information, have the most ex-situ occurrences. When positioned ex-situ, discourse-new information in the corpus is equally likely to occur sentence-initially and sentence-finally. A number of studies on discourse-driven word order variability in Russian (e.g, Neeleman and Titov 2009; Jasinskaja 2013; Sloussar 2010, 2011) draw attention to qualitatively different processes in discourse that call for fronting or post-posing an argument. Traditionally, in Russian, the sentence-initial position is associated with topics and contrastively focused arguments (Ionin 2001; Neeleman and Titov 2009), whereas non-contrastive new information is considered to favor the sentence-final, nuclear pitch accented position. Because of a limited number of contrastive foci in our materials, we did not explicitly test the effects of contrast or emphasis on argument linearization and leave this question to future research. The apparent interrelatedness of referent information status and ex-situ position confirms that word order variability in Russian is discourse-motivated.

Our results lend support to previous work on the interrelatedness of prosody and word order variability in Russian (Svetozarova 1998; Jasinskaja 2013) and select other free word order languages. According to our first hypothesis, a non-canonical word order in Russian has its own prosodic signature. Analysis of the prosodic expression of ex-situ words in our corpus confirmed this hypothesis. Fronting or post-posing a word relative to its canonical position affects its prosodic qualities via scaling (expanding or compressing) the magnitude of the prosodic parameters (f0 range, duration, and mean intensity) associated with that word. We presented evidence that acoustic-prosodic reflexes of an ex-situ position may be co-incidental with, and orthogonal to,

patterns of acoustic-prosodic variation indicative of the semantic, grammatical, and information-structural properties of a word. To illustrate, a post-posed ex-situ position in our data triggers reduction in f0 excursion size. However, f0 range undergoes expansion when the reader encounters an ex-situ post-posed word that introduces a conceptually salient animate referent that is discourse-new. Similarly, a fronted ex-situ position, often associated with topical, discourse-given information in Russian, receives a boost in intensity and a more dramatic expansion of f0 range, possibly due to the proximity to the leftmost boundary of a prosodic domain (Cho 2016). This effect, too, is not observed when the fronted word is easily inferable from context (r-bridging).

A qualitatively different source of acoustic-prosodic variation which we examined in this study is related to semantic and grammatical features of discourse referents. We formulated our second hypothesis to test if these referent-related features trigger parallel prosodic encoding in discourse. Our results show that in the read production data, the acoustic-prosodic measures of f0 range and mean intensity reached their maximum values for animate subjects, which are conceptually more salient than inanimate referents with object status. Vowel duration, for the majority of our speakers, showed systematic variation in response to referent animacy and was greater for animate referents. These findings are largely consistent with results of Antão et al. (2013) who examined prosodic variation related to subjecthood effects in Brazilian Portuguese. Antão et al. found that compared to objects, grammatical subjects have greater duration and f0 range, and that both measures are yet greater for utterance-initial subjects compared to non-initial subjects. One plausible generalization of these findings may be that subject-first effects, in part, are motivated not only in the propensity of grammatical subjects to occupy a position high up in the syntactic tree (Branigan et al. 2008), but also a position that is perceptually highly prominent in a sentence or phrase, due to being associated with default greater prosodic prominence.

Studies of the effect of animacy on grammatical function assignment and argument linearization (Prat-Sala 1997; Branigan and Feleiki 1999; van Nice and Dietrich 2003; Verhoeven 2014; Tanaka et al. 2005) report that cross-linguistically, animate entities, via greater accessibility and salience in discourse, tend to occur in higher syntactic positions as well as early in a sentence or clause. In this way, animacy has been proposed as a leading factor that determines grammatical function assignment and, in free word order languages, argument linearization. Our findings lend support to the special status of animate entities, in particular, grammatical subjects, via augmented prosodic expression. However, despite the observed variation in acoustic-prosodic parameters, no meaningful association was observed between referent animacy and its sentence position. Our results fail to support the finding of Lobanova (2011) who reported that sentence-initial nouns have a strong tendency to be animate, regardless of their grammatical function, in a corpus of 600 Russian sentences. In our materials, sentence-initial nouns were equally likely to be inanimate, moreover, slightly more inanimate nouns occurred sentence-initially. We attribute the lack of consensus with Lobanova's findings to the differences in study materials: our corpus is considerably smaller than Lobanova's and the distribution of different constituent orders in our materials is highly unbalanced.

One focus of the referent-centered analyses in this study is the effect of information status on acoustic-prosodic expression. In the read productions of the narratives, two information status categories, discourse-given and discourse-new, showed greater prosodic augmentation than discourse-bridging information. Whereas mean intensity was systematically greater for discourse-new information, in the reading performance of most of our participants, discourse-given information was associated with greater f0 range and duration than any other information category. One possible explanation for this finding may lie in the greater salience of discourse-given

referents, which have multiple mentions in discourse. This explanation, however, goes against the copious evidence that discourse-given information, which is more accessible due to its increased activation in the speaker's/reader's memory, is typically prosodically reduced (Aylett and Turk 2004; Watson 2010; Breen et al. 2010; Baumann and Riester 2013). We therefore hypothesize that the unexpected acoustic-prosodic effects of discourse-given status found in this study may stem from the reading mode of the production data or may reflect the ways in which individual speakers reconcile two distinct sources of prosodic variation, word order-based and referent-based. When crossed with other factors of interest investigated in this study, discourse-new status was in fact associated with augmented prosodic expression. Specifically, acoustic-prosodic parameters duration, mean intensity and f0 range were all greater for discourse-new post-posed animate subjects, relative to all other word types. Similarly, discourse new fronted objects had greater mean intensity and duration when the referent was animate and greater mean intensity when the referent was inanimate.

Overall, our findings present evidence for the incremental nature of acoustic-prosodic effects in Russian: concurrent prosodic processes that occur during speech production originate in qualitatively distinct discourse phenomena. At a more foundational level, the invariant semantic feature animacy and the related grammatical feature subjecthood scale prosodic parameters such that greater prosodic prominence is given to referents that are animate and grammatical subjects. As she progresses though discourse, the reader perceives information as new, given, or easily inferable. At the level of information structure, each information status is associated with distinct positional and acoustic-prosodic properties. In the materials used in this study, discourse-new information is a highly 'mobile' and prosodically prominent information category. In individual speakers' productions, the prosodic expression of new and given information was often influenced

by the ordering of sentence constituents. This leads us to conclude that referent information status in discourse may be signaled by placing the critical argument into a designated position in a sentence or clause, where it may attract greater attention of the reader or hearer, by virtue of being structurally and prosodically distinct.

Given considerable disagreement in the literature on the syntactic derivation of non-canonical orders in Russian (e.g., Slioussar 2011b, Baylin 2004, 2005), we remain uncommitted as to the syntactic structures associated with ex-situ positions in Russian. We also make no claims about prosodic phrasing associated with non-canonical word orders. The findings of this study invite a straightforward and falsifiable proposal that referent status and information structural constraints have an impact on argument linearization in Russian resulting in prosodic variability that is mediated by structural variability in discourse. In our ongoing work, we evaluate this proposal by testing perceptual processing of ex-situ words in discourse (Luchkina and Cole in preparation).

Speaking mode effect and inter-speaker variability

One important corollary we add here concerns the possible influence of the speech mode on the results of acoustic-prosodic analyses reported in this study. Read speech has been characterized as syntactically more complex than spontaneous speech and may be articulated more slowly and with greater pitch range and pitch declination (Swerts, Strangert, and Heldner 1996). The use of read speech has been shown to affect results of prosodic analyses, as reported in Baumann and Riester (2013) and de Ruiter (2015). Baumann and Riester (2013) tested empirical applications of RefLex information structure scheme using two corpora of German speech, spontaneous and written. They reported that prosodic encoding of information status was most apparent in their read speech corpus, where pitch accent distribution among information categories was highly consistent with the expectation that novel information in discourse is accented, whereas given information is

deaccented. De Ruiter's (2015) analysis of German, similarly, shows that the rate of deaccenting of discourse-given referents is greater in read speech. De Ruiter attributes this effect to the more careful use of prosody and decreased cognitive load during reading than during spontaneous speech. The only study of accentuation patterns reflecting referent information status in Russian that we are aware of was conducted by Sityaev (2000). Sityaev used a corpus of read speech to compare accenting of discourse-new and discourse-given referents. Contrary to results presented by Baumann and Riester (2013) and de Ruiter (2015) for German, Sityaev reported that an unusually high proportion (79-97%) of discourse-given referents in his corpus of read Russian speech were accented.[11]

In the current study, the special status of reading intonation may be slightly mitigated due to the selection of study materials. Recall that one of the narratives we used was a fable, highly conversational in nature and with multiple instances of direct speech. The second narrative, however, was an excerpt from a biography and featured characteristics more typical of written language, including longer sentences and a stricter adherence to standard language register. Text genre and read speech mode therefore may have affected the production styles of the speakers, leading some to use prosody as a cue for sentence structure more than for referent features. Visual examination of the predicted marginal effects plots using individual speakers' production data revealed that prosodic encoding of animacy, subjecthood, and information status of a word often interacted with the position of that word in a sentence or phrase (see Figures 2.3.-2.5. and 2.8.-2.10.). Specifically, the finding that fronted discourse-given words featured increased f0 range in most speakers' production, whereas those occurring in-situ or post-posed did not, suggests that

[11] We treat Sityaev's (2000) results as preliminary. The author used a dichotomous accenting scheme, with levels 'accented' and 'unaccented'. H* notation was applied to all instances of accented words without further commentary on which criteria or reliability analyses were used when labeling accents.

such outstanding prosodic marking may be conditional on word order. Our preliminary analyses of speaker-specific differences in prosodic encoding calls for future work exploring this issue. In particular, the question of which cognitive features, oral production styles, or narrative-specific properties account for what we refer to as a 'high' or a 'low' profile speaker merits future investigation.

Conclusion

As we speak, prosody signals information about discourse entities and their grammatical relationships and information status in discourse. Results of this study reveal prosodic effects due to the animacy of discourse referents and their grammatical function, and due to dynamic information status grounded in the knowledge state of the speaker or reader. In a free word order language like Russian, prosody is not the only means of encoding discourse information. Results demonstrate that depending on the status of the information that a word contributes to discourse, it may be more or less likely to occur in an ex-situ position in a sentence or phrase. We presented evidence that sentence-initial and sentence-final ex-situ positions have their own prosodic signatures which may be orthogonal to the information value of a word or its relative discourse salience. Preliminary examination of individual speakers' read productions has revealed that the two underlying sources of prosodic encoding, structural and referent-based, come together in a variety of ways, frequently reflecting the dominance of one of these sources over the other in their interaction.

CHAPTER 3

PROSODIC AND STRUCTURAL CORRELATES OF PERCEIVED PROMINENCE IN RUSSIAN

(Paper 2)

Abstract

Linearization of sentence constituents in Russian, a highly free word order language, depends on the information structure in discourse. A change in word order reflects the ex-situ word information status, and affects that word's acoustic-prosodic expression. The acoustic-prosodic changes that the ex-situ words undergo in the read production by a female native speaker include augmentation of acoustic-prosodic parameters intensity, duration, and fundamental frequency associated with the ex-situ fronted words and reduction of intensity and pitch range in otherwise lengthened ex-situ post-posed words. In two unguided prominence rating tasks, Russian speakers rated perceived prominence of words in written and spoken discourse. Prosodically prominent words and words occurring in non-canonical positions were more likely to be rated as prominent. Results reveal that typologically distinct cues to perceived prominence are deployed in Russian. They include acoustic-prosodic augmentation of the prominent word and positioning of the prominent word at a phrasal boundary. Listeners use acoustic-prosodic expression of ex-situ words in a probabilistic fashion, whenever the magnitude of a prosodic parameter is inversely related to that word's accessibility in discourse. Although prosodic cues explain more variation in the prominence scores than the word order variability, a combination of the word position relative to the canonical constituent order and that word's prosodic expression yield optimal modeling of perceived prominence in Russian.

Introduction

An essential aspect of comprehending language, in written or spoken modalities, is interpreting the status of a linguistic entity relative to the discourse or narrative context. A word or phrase can be introduced as information that is new to the discourse and as relatively important or emphasized, or as already known information of lesser significance. Cross-linguistically, prosody and linear organization by which we mean linear order of presentation of information in sentence or phrase, provide some well-known means of expressing word-level prominence (Morgan, Meyer, and Newport 1987, Swerts, Krahmer and Avesani 2002). Languages may deploy one of these prominence cues, to the exclusion of the other, or use a combination of word order and prosodic mechanisms to express prominence. To illustrate, in English, prominence marking by means of prosodic cues is particularly important, since the room for structural reorganization of an utterance in discourse is limited. Multiple studies have shown that prominent information in English, often identified as sentence focus, undergoes prosodic augmentation, as evident from greater measures of acoustic-prosodic parameters intensity, segment duration and fundamental frequency (f0) associated with the focused word (Xu and Xu 2005, Kochanski, Grabe, Coleman, Rosner 2005, Watson 2010). Structural prominence in English is possible, however, very limited (Arnold, Losongco, Wasow, and Ginstrom 2000). Cleft constructions of the kind '*It's Mary who kissed John'*, where the focused word, '*Mary*' is structurally (and inonationally) prominent, are reported to have stylistic coloring (Skopeteas and Fanselow 2008) and may require greater cognitive effort during processing due to a more complex syntactic derivation (Ferreira 2003, Gordon, Hendrick, Johnson 2001). Rather than structurally reorganizing an utterance, English speakers demonstrate preference to highlighting the prominent word by means of intonation, as in '*MARY kissed John*'.

In French, unlike in English, the main prosodic prominence is invariably located at the rightmost boundary of a phonological phrase, which often coincides with the sentence-final word (Féry 2001, 2003). Féry argues that as a non-lexical stress language, French deploys prosodic cues to signal prosodic phrasing only. As a result, in French, structural reorganization of an utterance, often resulting in a cleft construction, presents a key way of expressing prominence. Consistent with this view, Skopeteas and Fanselow (2010) observed a high proportion (33%) of clefts used to mark a focused subject in a corpus of 297 semi-elicited utterances in Québec French. In the same study, the rate of clef constructions observed in a corpus of 118 semi-elicited utterances in English, was as low as 5%.

While English and French are examples of languages which strictly adhere to one prominence marking strategy, prosodic or structural, a number of languages, including Hungarian, Georgian, Finnish, Romani, and Russian, have been reported to deploy both these cues, interchangeably or in combination (Szendrői 2001, 2003, Genzel, Ishihara, and Surányi 2015, Skopeteas and Fanselow 2010, Vainio and Järvikivi 2006, Sekerina and Trueswell 2012). One fundamental difference between these languages and a language like English or French is that they allow surface reorganization or, simply put, reordering of sentence constituents which does not involve clefts or higher order syntactic derivations.

Consider the following example from Komatini Romani:

(3.1) Me ˈapora peraˈdom
my pills lost.1SG
I lost my PILLS.” (Arvaniti and Adamou 2011)

Arvaniti and Adamou (2011) report that the canonical constituent order for Romani is VO. By positioning the direct object pre-verbally, as shown in the example (3.1), the speaker brings it into focus and renders it prominent.

In free word order languages like Komatini Romani, constituent reordering is usually context-dependent and may be motivated by a number of factors which are closely related to prominence. One of these factors is the tendency to distribute information in a sentence or phrase in an optimal way. Such optimal distribution of information often calls for information already known to the speaker and the listener to precede the information which is new. The optimal distribution of information, therefore, assumes a linear precedence of topics (given information) before foci (new information) (Kiss 1998, Arnold et al. 2000, Clifton and Frazier 2004). Non-canonical utterances form, whenever adherence to the canonical order would compromise the preferred information structure by positioning discourse-new information before discourse-given information. Such outcome would be suboptimal for processing and comprehension purposes.

Language-specific prosodic constraints may present another motivation for constituent reordering. According to Kiss (1998) and Szendrői (2001, 2003), in Hungarian, canonically a VSO language, positioning a constituent pre-verbally is required for that constituent to receive a distinctive prosodic manifestation and be interpreted as (contrastively) focused. Similarly, in right-branching free word order languages, including Russian and Italian, moving a focal constituent to the sentence-final position aligns it with the phrasal nuclear pitch accent and confers its greater perceived prominence (Sloussar 2010, 2011a, Calhoun 2015, Féry (2013).

Donati and Nespor (2003) and Swerts et al. (2002) proposed a two-way classification of languages depending on whether their grammar permits to express focus in situ, by varying the location of prosodic prominence in the sentence and thereby preserving the canonical order of sentence

constituents. Donati and Nespor (2003) proposed the term 'prominence dislocating' for languages which use prosodic cues to mark the prominent focused constituent. In these languages, the location of the main prosodic prominence in the utterance is variable and aligns with the focused word. Reflecting on the variable location of prosodic prominence, Swerts et al. (2002) referred to prominence-dislocating languages as 'plastic'. Donati and Nespor proposed the term 'constituent dislocating' for languages with demonstrated preference for constituent reordering as a focus cue (c.f. Swerts et al. (2002) who used the term 'non-plastic'). Past the differences in the terminology, Donati and Nespor (2003) and Swerts et al. (2002) support this two-way classification of languages with data from strictly constituent-dislocating languages like Italian and strictly prominence-dislocating languages like Dutch and English. Other studies, however, draw attention to a more complex prominence marking strategy, whereby a speaker alters the order of sentence constituents *and* (probabilistically) deploys prosodic variability to signal prominence. Returning to the example (3.1.), Arvaniti and Adamou (2011) report that the reversal of subject and object linear order in Romani is combined with deaccenting of the verb 'lost' and prosodic augmentation of the focused and highly prominent constituent 'pills'. Similarly, Patil, Kentner, Gollrad, Kügler, Féry, and Vasishth (2008) report that prosodic expression of ex-situ constituents in Hindi, an SOV language, differs from that of their in-situ counterparts. Patil et al. documented that under a non-canonical OSV order in Hindi, the ex-situ post-posed subject resists post-focal compression when the sentence-initial object is in focus. Additionally, the fronted (discourse-given) object undergoes prosodic strengthening and features greater pitch range and maxima even when the post-posed subject is in focus. Further evidence corroborating the observation that a change in constituent ordering may be accompanied with acoustic-prosodic modifications in the expression of the ex-situ constituent has been presented for Hungarian (Genzel et al. 2015), Russian (Luchkina and

Cole 2016), Greek (Botinis et al. 2005), Samoan (Calhoun 2015) and other free word order languages.

While the evidence that languages may combine constituent reordering and prosodic accenting to express focus and, more generally, prominence, continues to accrue, less is known about the relative functional significance of these prominence marking mechanisms during perception of spoken language. Two interesting questions arise with this respect. The first of these questions concerns the distribution of prosodic and word order variability and its use as cues to *perceived* prominence in a free word order language. The second question concerns the relative importance of prosodic and word order cues for perception in the contexts where these cues are simultaneously available.

To answer these questions, in this work, we focus on the concurrent application and functional significance of prosodic and word order variability to perceived prominence in Russian, a case marking free word order language. We seek to determine how prosodic and word order variability is utilized during production and comprehension of discourse, and gauge the relative contribution of acoustic-prosodic variation and word order variation to perceived prominence when these cues are used in combination. In order to understand which factors guide naïve readers' or listeners' perception of a word as prominent in a discourse or narrative, we conduct a prominence rating experiment with linguistically naïve native Russian speakers. We parameterize word-level perceived prominence ratings with respect to acoustic-prosodic features and word order features of discourse and gauge their relative functional validity in Russian. Russian is chosen as the test case because it allows, but does not require reordering of sentential constituents for information structure purposes and exhibits distinctions in prosodic prominence among the sentence constituents (Sekerina 2003, Slioussar 2011a, Sekerina 1999).

This paper is organized as follows. We begin by discussing the notions of relative [information] accessibility, information status and the related notion of perceived prominence. Next, we summarize what word order and acoustic-prosodic variability are manifest in Russian. We examine read production of two published narratives by a native Russian speaker and discuss how prosodic and word order variation work separately and in concert as cues to information status, a proxy of word-level perceived prominence. Following a presentation of the analyses of the read production data used in this study, we present the methodology and results of two unguided prominence rating tasks, in which linguistically-naïve native Russian speakers provided word-level perceived prominence ratings while reading or listening to two stylistically distinct narratives. Analyses of perceived prominence ratings reveal that albeit simultaneously available, prosodic and word order cues in Russian are not fully functionally homogenous: while both cue types contribute to the likelihood that a word is perceived as prominent, acoustic-prosodic expression of ex-situ words may be compromised by the proximity of a word to a phrasal boundary. Next, we discuss these results, and what they mean for our understanding of multi-cue environments such as free word order discourse.

What is perceived prominence

Information accessibility and the related notion of information prominence have been offered a variety of interpretations in linguistic and psychological literature (e.g., Watson 2010, Arnold et al. 2000). In cognitive accounts (Chafe 1976, 1994, Lambrecht 1994), accessibility or givenness of discourse entities is described in terms of the activation costs associated with brining these entities into the focus of speaker's/hearer's attention. Information accessibility may be viewed as gradual or continuous, categorically discretized in production through lexical choices, and prominence marking devices, such as prosody, word order, or morphological markers (Morgan et

al. 1987). One way to conceptualize accessibility is through a number of discrete linguistic categories in discourse which reflect the information status of words and the extra-linguistic referents identified by these words. One of the most intuitive dichotomies in this sense is that between given, more accessible, shared information and novel information, which is less accessible and calls for an update in the mental state of the hearer to reflect the new knowledge communicated by the speaker (e.g., Krifka 2007).

Capturing varying degrees of information accessibility, Chafe (1976) and Prince (1981) adopted a tripartite distinction of discourse entities into (1) the most accessible or given, active, (2) least accessible or novel, inactive and (3) inferable, semi-active. Givennes and newness in discourse may be further specified at more than one level. Baumann and Riester (2012, 2013) used two corpora of German text, read and spoken, and evaluated the information status of discourse entities at two distinct levels of representation, lexical and referential. The lexical level of representation pertains to the choice of words previously mentioned or novel, regardless of the extra-linguistic referents identified by these words. The referential level of representation pertains to the newness of discourse referents, regardless of the specific lexical items used by the speaker to identify these referents. Consistent with this dichotomy, at the referential level, the information status reflects whether a referent 'has occurred before', which makes it *referentially-given* (henceforth, r-given) or is introduced for the first time, which makes it *referentially-new* (henceforth, r-new). At the lexical level, the information status reflects whether a lexical word occurs in discourse repeatedly, which makes it *lexically-given* (henceforth, l-given), or for the first time, which makes it *lexically-new* (henceforth, l-new). For illustration purposes, consider the following example from Baumann and Riester (2013, p.12):

(3.2) A car was waiting in front of the hotel. I could see a woman in the car.

r-new r-given

l-new l-given

In (3.2), we observe a progression of the discourse entity ‘car’ from the r-new status (first mention of a referent) to the r-given status; we also observe a progression of the lexical word ‘car’ from the l-new status (first mention of a lexical word) to the ‘l-given’ status (repeated mention of a lexical word).

Prominence is a multi-faceted phenomenon and can be defined with reference to various inherent syntactic and semantic features of discourse referents and lexical items which identify those referents. At the syntactic level, such features include grammatical function and linear order of arguments in discourse; at the semantic-pragmatic level, they include, but are not limited to, information status, thematic relations, specificity, and animacy. Some known prominence scales rank discourse-novel referents above discourse-given referents, specifics above generics, animates above inanimate entities, thematic agents above thematic patients (Branigan, Pickering, Tanaka 2008, Arnold et al. 2000, Verhoeven 2009, 2014). During discourse comprehension, the relative prominence of discourse entities conditional on their inherent and context-dependent features informs their *perceived prominence* and determines how prominent they appear to a linguistically naïve reader or listener.

Bornkessel-Schlesewsky and Schlesewsky (2009) argue that relative information prominence plays a crucial role in language comprehension and aids establishment of interpretative relations in discourse. During written or spoken language use, it is often disambiguated through the use of special prominence marking mechanisms, such as prosodic strengthening and/or accenting, a morphological prominence marker, or positional prominence marker (Féry 2013). Speakers and

listeners use these cues to form personal discourse experience leading to perception of information at varying degrees of prominence. Perceived prominence is a product of more than one cue. It is richly encoded by means of multiple distinct channels which relay lexical, syntactic, prosodic and other kinds of information about the relative newness, accessibility, or salience of discourse entities (Morgan et al. 1987, McKoon, Ratcliff and Ward 1993, Watson 2010). To illustrate, in a study by Cole, Mo, and Hasegawa-Johnson (2011), English listeners rated words as prominent based on their acoustic-prosodic expression, which the authors characterize as a signal-driven correlate of prominence, lexical properties (e.g., lexical frequency), and listeners' previous experience with discourse (e.g., number of previous mentions), characterized as expectation-driven correlates of prominence. While acknowledging the inherent complexity of perceived prominence, in the following investigation, we limit our attention to two types of its correlates, acoustic-prosodic and structural, achieved via a particular linearization of the sentence constituents.

Prosodic variability as a prominence marking mechanism

The use of prosody as a prominence cue presents a well-understood psychologically real property of spoken language use (Watson, Arnold, and Tanenhouse 2008, Mo, Cole, and Lee 2008, Watson 2010, Bishop 2012, Bolinger, 1986, Pierrehumbert & Hirschberg, 1990, Breen, Fedorenko, Wagner and Gibson 2010). It involves perceptually salient changes in voice quality, length and loudness, the psycho-acoustic correlates of duration and intensity, as well as changes in the perceptual correlate of f0, pitch. Acoustic-prosodic strengthening or accenting of information has been characterized as a source of important cognitive benefits during discourse comprehension. A number of psycholinguistic studies have shown that nuclear pitch accented words enhance sentence processing and are processed efficiently (Cutler, Dahan, Van Donselaar 1997). More

recently, Fraundorf et al. (2010) showed that prosodically-expressive speech promotes better understanding and subsequent recall of information. Fraundorf et al. tested recognition of contrastive referents by English speaking listeners 30 minutes after they had been presented in short stories. They reported that the odds of correct target recognition were 1.5 times better when the auditorily presented target was H* accented and 3.4 times better when the auditorily presented target was L+H* accented. These results confirm that prosodically prominent words are perceptually more salient than the non-prominent words.

In a fixed word order language like English, it is generally the case that a word can be assigned prosodic prominence as an expression of its information status and relative accessibility in discourse regardless of its position in the utterance (Gussenhoven 2007, Ladd 2008, Swerts et al. 2002). Calhoun (2010) argues that relative 'informativeness of a word' affects its prosodic expression and determines its alignment with the metrical (prosodic) structure of an utterance. Hence, the most informative element in a sentence or phrase aligns with the prominence-lending site and is perceptibly different from the rest of the utterance. Given the great saliency of prosodic cues in expressing information status and relative information prominence, Calhoun (2010) proposed that higher-level post-lexical prosodic structure of a language is based on the information structure.

Jun (2014) further distinguished between two types of languages which utilize prosodic distinctions as prominence cues. Jun's proposed classification is based on which element of the prosodic structure of a language is used to express prominence. According to this classification, 'head prominence languages' are those which accomplish phrase-level prominence marking on the phrase head. In these languages, a prominence-lending pitch movement, in the form of pitch accents, tones, or lexical stress, is aligned with the head of a prosodic domain. 'Edge prominence

languages', on the other hand, mark prominence by aligning the prominent word with a phrasal boundary, where it is marked with a boundary tone.

As an illustration, consider how prominence is expressed in English, a right-branching head prominence language (following Jun's classification). Sentence-final prominence-lending pitch movement, or nuclear accent, by default, aligns with the rightmost sentence constituent in English, as shown in (3.3)B. Under a discourse configuration which lends a non-sentence-final word prominent, the mechanism of metrical reversal or stress shift (Calhoun 2010, Neeleman and Reinhart 1998) displaces phrasal prominence leftward and aligns it with the prominent constituent, as shown in (3.4)B:

(3.3) A: What happened to all the garbage?

B: The garbage was taken OUT.

(3.4) A: What happened to this room?

B: The GARBAGE was taken out.[12]

The categorical distinction between prosodically 'prominent' vs. 'non-prominent' elements rests on the gradient differences in the acoustic-prosodic parameters which contribute to the prosodic expression at a word level and phrasal level. Many languages utilize similar acoustic-prosodic cues to express the information status of a word and its perceived prominence. Breen et al. (2010) report that in English, the location, size, and type of the prominent constituent is determined based on the relative magnitude of its acoustic-prosodic features intensity, duration, and pitch. Xu and Xu (2005) find that narrowly focused words in English have consistently higher pitch peaks combined

[12] Examples from Calhoun (2010, p.15). Words typed in upper case letter are focused.

with magnified stressed vowel duration; the post-focus region, on the contrary, is characterized by compressed pitch range. Dutch is another language which heavily relies on prosodic cues for signaling prominence. Hanssen, Peters, and Gussenhoven (2008) report that in Dutch, duration of the nuclear accented syllable, as well as timing, magnitude, and slope of pitch peaks are affected when the speaker conveys novel information or narrow focus. The observation that prosodic accenting and de-accenting successfully reflect relevant information properties, such as accessibility and givenness/newness holds, cross-linguistically (Ladd 2008, Cruttenden 2006, Jasinskaja 2013, Breen et al. 2010, Watson 2010).

Watson (2008) and Aylett and Turk (2004) proposed that prosodic prominence is inversely related to referent accessibility in discourse: less accessible referents are more likely to be prosodically prominent, whereas more accessible referents are more likely to be prosodically reduced. Consistent with this proposal, in a corpus of read German speech, Baumann and Riester (2013) found that lexically- and referentially-new nominal expressions were systematically pitch accented[13], whereas referentially and lexically given items appeared de-accented.

Word order variability as a prominence marking mechanism

Languages have a number of constituent linearization possibilities at their disposal, constrained by their syntactic and phonological properties (Féry 2013). Under relatively free word order, linearization of sentence constituents reflects their information status and relative prominence.

In a series of reading and probe recognition experiments, McKoon et al. (1993) tested the importance of the constituent position in a sentence during discourse comprehension in English.

[13] The most prosodically prominent category of information in Baumann and Riester's study is associated with nominal expressions which are both lexically- and referentially-novel, in that they introduce previously unmentioned discourse referents and are expressed by previously unused lexical items (Baumann and Riester 2012).

The authors main argument is that the 'surface form of an utterance' reflects the relative status of discourse concepts by assigning different degrees of prominence to different syntactic positions. McKoon et al. (1993) found that in English, placing a constituent in syntactically prominent positions, such as a sentence-initial subject position, increases its accessibility and may facilitate its retrieval from long- and short-term memory. Bock and Warren (1985) proposed that higher grammatical functions, in particular those of sentence subjects, tend to be realized relatively early in the sentence due to their greater conceptual accessibility and easier retrieval from memory. Conceptual accessibility of discourse entities is closely related to their information status in discourse. By assigning more accessible, previously mentioned information to higher sentence positions, the speaker maintains the precedence of discourse-given information before discourse-new information and possibly facilitates processing and comprehension for the listener or reader (Arnold et al. 2000, Branigan et al. 2007).

Particularly compelling evidence for the importance of argument linearization in a sentence or phrase comes from 'free word order' languages, referred to as 'constituent-dislocating' by Donati and Nespor (2003). In such languages, grammatical function of sentence constituents is expressed by means of word morphology, irrespective of where in the sentence a word occurs. As a result, the speaker may alter the surface order of sentence constituents for pragmatic purposes. In most free word order languages, there exists a canonical ordering of sentence constituents which is most common and is strongly preferred across discourse contexts. By altering this canonical order, speakers put constituents into designated sentence positons, where they are more (or less) likely to be perceived as prominent. The mechanism of constituent displacement, whereby a word appears ex-situ relative to the expected, canonical position, enables the speaker to express information prominence and bring the critical information to the focus of the reader's or listener's attention.

To illustrate, in Italian, a relatively free word order language, the ordering of sentence constituents can cue its information status (Donati and Nespor 2003). The canonical word order in Italian sentence is SVO, as shown in (3.5).

(3.5) Mario e`arrivato.

Mario is arrived

'Mario arrived'

The reverse ordering of S and V, less frequent than the canonical SV order, confers greater prominence to the discourse-new subject, as show in (3.6):

(3.6) E' arrivatoMario.

Is arrived Mariob.

'Mario arrived' (examples from Donati and Nespor 2003, p.1137)

Donati and Nespor (2003), Féry (2013), and Genzel et al. (2015) argue that the unambiguous marking of the focused constituent by means of word order obviates the special use of prosody, whereby a prosodic accent aligns with the prominent focused constituent. However, the role of prosody in non-canonical strings is far from trivial. By altering word order, the speaker preserves 'default sentence prosody' (Genzel et al. 2015) and utilizes a prosodically strong position at the phrasal level to signal prominence. Such prosodically strong position may be determibed by the default location of the nuclear pitch accent or a prosodic domain boundary. Preservation of the default prosodic structure by means of constituent reordering has been characterized as one possible motivation behind word order 'flexibility' (Szendrői 2003, Calhoun 2010, Féry 2013). To illustrate, in Hungarian, a language where word order must alter to mark narrow focus, the contrastively focused constituent obligatorily moves to the pre-verbal position, where it receives

the nuclear phrasal pitch accent (Szendrői 2003). As a result, the narrowly focused word is not only structurally prominent in Hungarian, but also occurs in the position which is prosodically prominent. Confirming this, Genzel et al. (2015) documented that ex-situ narrowly focused words in Hungarian have distinctive pitch scaling and contours.

Calhoun (2010, 2015) proposed that reordering of sentence constituents reflects adherence to the prosodic properties or 'prosodic requirements' of a language, whereby the speaker (obligatorily or optionally) moves the focused word to align it with the position of maximal prominence, such as the nuclear accent at the phrasal level. Calhoun (2015) supported this proposal with evidence from Samoan, an Austronesian VSO language. In Samoan, the sentence-initial, pre-verbal position is maximally prominent, both from the structural perspective and in terms of its prosodic expression. Focus fronting affects the prosodic realization of the focused constituent, by making it maximally prominent and deaccenting the post-focal material. Using this evidence, Calhoun argues that Samoan demonstrates prosodically motivated syntactic movement as a focus marking strategy. In a similar vein, Féry (2013) argued that word order variability, cross-linguistically, serves as a mechanism permitting the speakers to mark prominence on the edges of prosodic domains, by aligning the prominent (focused) element with a phrasal boundary. Such mechanism operates under the assumption that alignment with a boundary of the largest available prosodic domain (such as a prosodic phrase or intonation phrase) serves as a universal focus cue. Via prosodic boundary or nuclear pitch accent site alignment, the focused (prominent) word may further receive a special prosodic expression, by sharing the acoustic-prosodic properties associated with the designated prominence-lending site in a given language. Féry (2013) further argues that the greater the prominence of the focused element, the more plausible it is to re-order sentence constituents so as to align the focused element with a prosodic boundary (and thereby mark it as prominent).

Similarly, foci which are relatively high on the prominence hierarchy, such as contrastive foci, are more likely to be prosodically prominent, even if the prominence marking alignment has been fulfilled.

Cross-linguistic evidence

The choice of the prominence marking strategy may be viewed as a syntactic structure preservation strategy, accomplished when the prominent constituent is cued in-situ or, alternatively, as a prosodic structure preservation, strategy, accomplished when the underlying prosodic structure remains unchanged but the canonical constituent order is compromised. Under this scenario, natural language deploys one preferred way of expressing prominence, via prosody or via constituent ordering. In support of this account, in a comparative study of Italian and Turkish (relatively free word order languages), and English (a rigid word order language), Donati and Nespor (2003) show that languages with rigid word order allow prominence to be signaled by means of prosody at different locations in the sentence. Languages with flexible word order, express prominence through constituent reordering and exhibit less variation in the location or magnitude of acoustic-prosodic parameters. This model, wherein a language uses one cue type, prosodic or word order, to express prominence, generates a prediction that speakers should disprefer to engage word order *and* prosodic marking interchangeably or simultaneously to cue prominence, possibly, for reasons of cognitive economy. Skopeteas and Fanselow (2011) argue that cross-linguistically, prosodic accenting presents the more economical way of signaling prominence, compared to constituent reorganization (cf. Féry (2013), who argues that constituent ordering relative to a prosodic domain boundaries is the universal (and primary) focus cue, probabilistically reinforced with (secondary) prosodic cues).

This is so because of the relatively simpler derivation of the canonical order, with the prominent

word marked in-situ The assumption that a change in word order and prosodic augmentation as prominence marking mechanisms are in complementary distribution in a given language is supported with evidence presented in the empirical study by Swerts et al. (2002). Swerts and colleagues conducted a production-perception study of accentuation in relation to discourse-new and contrastive information in Dutch, a language with limited word order variability, and Italian, a language with free word order. Consistent with Donati and Nespor's proposal, Swerts et al. found that prosodic marking of prominent discourse-novel and contrastively focused information was systematically deployed by Dutch speakers and listeners. Accentuation patterns found in the Italian data, however, did not relate to the either of these information categories. Since production data elicitation in Swerts et al.'s study was experimentally constrained such as to discourage (Italian) speakers to alter word order, no evidence for use of of constituent linearization as a prominence cue in Italian was gathered.

Work by Swerts et al. (2002) and Donati and Nespor (2003) provides evidence in support of the two-way classification of languages into prominence-dislocating or plastic vs. constituent-dislocating or non-plastic. In this work, however, we depart from the assumption that prosodic only *or* word order only mechanisms are engaged to express prominence. We examine a different scenario, whereby word order *and* prosody are used by speakers as interchangeable or simultaneously deployed prominence marking mechanisms. Earlier studies indicate this possibility for a number of free word order languages, including Romani (Arvaniti and Adamou 2011), Greek (Baltazani 2003), Georgian (Skopeteas and Fanselow (2010), Finnish (Vainio and Järvikivi 2006), and Hindi (Patilet al. 2008). In these languages, prominence can be cued via strategic positioning of a word relative to other words in a sentence or phrase and via prosodic marking, because word order variability and prosodic variability are concurrent and happen simultaneously.

To date, rather few attempts to analyze the cross-application of prosodic and linearization cues in spoken language use have been presented. Georgian is one language, in which prosodic augmentation and reordering of phrasal constituents may be used interchangeably, or, sometimes, in combination. Skopeteas, Féry, and Asatiani (2008) and Skopeteas and Fanselow (2010) reported that in Georgian, prominence may be signaled in-situ, by means of prosody, and ex-situ, by means of word order. The two prominence-marking strategies map onto two different foci types, in-situ (new) information focus and ex-situ contrastive focus. Probabilistically, ex-situ prominent words in Georgian also feature a distinctive pitch contour, due to an increase in pitch height (Skopeteas et al. 2008). Similarly, Patil et al. (2008) found that in Hindi, canonically an SOV language, pitch excursion and segment duration are augmented when the object occurs pre-verbally in the non-canonical OSV order, where it is contrastively focused. Arvaniti and Adamou (2011) reported new information foci in Komotini Romani are signalled via a combination of pitch accenting and constituent reordering employed in the same utterance. Evidence that the prominent constituent appears ex-situ and receives a distinctive prosodic expression has been also presented for Greek (Baltazani 2003, Botinis et al. 2005), Finnish (Vainio and Järvikivi (2006), and Russian (Botinis et al. 2005, Luchkina and Cole 2016) and will be reviewed in more detail in the following sections.

To summarize, we have shown that prosodic and word order variability are widely used, cross-linguistically, to express prominence. The application of these mechanisms is constrained not only by the discourse properties, such as information structure, but also by the syntactic and prosodic characteristics of a language. In the languages in which prosodic cues and word order may be simultaneously engaged in expressing prominence, a change in the linear order of sentence constituents alters the pragmatic interpretation of the sentence and affects the acoustic-prosodic expression of the ex-situ word.

For languages which may concurrently apply prosodic cue and a change in word order to mark prominence, two relevant empirical questions arise:

(i) What is the status and the distribution of word order and prosodic variability in discourse, at the word level, and how do they interact?

(ii) In the case when constituent linearization and prosodic variability jointly act as prominence cues, are they functionally equivalent?

In this study, we examine these questions for Russian.

Prosodic prominence and structural prominence in Russian

Word order.[14] Russian is a morphologically rich language with overt case markers. The semantically neutral, default word order in Russian is SVO. As in other free word order languages, words in a sentence can appear in-situ, fronted, or post-posed, relative to their canonical positions. Linearization of sentence constituents serves information structure purposes via marking information status or communicative intent and not grammatical function (Bryzgunova 1980, Yokoyama 1986, Neelman and Titov 2009, Slioussar 2011a, Svedova 1982, Sekerina 2003).

Jasinskaja (2013) proposed that in Russian, via the mechanism of *word order optimization*, the preferred location for discourse-given information is at the left edge of a sentence or phrase, and the preferred location for discourse-novel information is sentence-final. Consistent with this configuration, new information foci in Russian (optionally) undergo movement to the sentence-

[14] Within generative grammar framework, non-canonical constituent orders in Russian are said to be derived via scrambling (Bailyn 2004, Ross 1967). This approach assumes that in non-canonical orders, constituents are scrambled from base-generated positions into landing positions and that only the surface (as apposed to base-generated) constituent order changes. In this paper, we adopt a different, functional approach to word order variability in Russian, offered in the work by Bryzgunova (1977), Yokoyama (1986) and much subsequent related work. Under the functional approach, word order variability is contextually constrained and is motivated by information structure properties of discourse.

final position, where they align with the nuclear pitch accent, whereas contrastively focused discourse-given topics may undergo optional fronting and occur pre-verbally (Neeleman and Titov, 2009). As a result, word order alterations are context constrained (see Chapter 1, example 1.8).

In the study reported in Chapter 2, Luchkina and Cole (2016), we have examined the relationship between information status and constituent ordering in Russian. We focused on the distribution of three categories of discourse referents, given, inferable (easily predictable in context), and novel, across in-situ, ex-situ fronted and ex-situ post-posed sentence positions in two published narratives in Russian We reported that referents which are new to discourse present the most 'mobile' information category in Russian and account for more ex-situ positions than any other category of information. In our materials, discourse-new referents, least accessible and high in information value Fisher and Tokura, 1995; Fowler and Housum, 1987; Chafe, 1994; Clark and Haviland, 1977; MacWhinney and Bates, 1978; Marslen-Wilson, Levy, and Tyler, 1982; Prince, 1981, Katz and Selkirk 2011), were particularly common in the post-posed (utterance-final) ex-situ position, which is obligatorily pitch-accented in Russian. Consistent with Jasinskaja's (2013) proposal, discourse-given and easily inferable referents had a tendency to occur in-situ or ex-situ, fronted to the left edge of a sentence or phrase. These results constitute confirmatory evidence that word order variability in Russian is engaged in expressing information structure in discourse. One logical possibility related to this conclusion is that speakers may assign a word to a designated position in the sentence to reflect, among other things, of its relative prominence in discourse.

Despite the fact that Russian is often characterized as a highly free word order language, the canonical SVO order, by far, presents the most common word order and accounts for approximately 80% of the utterances in the Russian language corpora (Bivon 1971). As a result,

80% of the time, Russian is necessarily 'reduced' to a fixed word order system. Consistent with this observation, canonical linearization of sentence constituents in Russian is compatible with different information structure configurations, which means that prosodic cues may be called upon in order to signal the location of prominent information in a canonical SVO sentence.

Acoustic-prosodic variability. Under the canonical constituent order, the nuclear pitch accent in Russian is realized inside the focused constituent (Jackendoff 1972). By default, the focused constituent is the sentence-final argument and presents 'the main locus of prominence' in the sentence (Meyer and Mleinek 2006).

Sekerina and Trueswell (2012) emphasize the important role of sentence prosody as a means of expressing information structure in Russian. In the earlier work which examined intonation correlates of information structure in Russian, discourse-new information was characterized by a distinctive falling pitch contour and greater pitch peak height[15] (Neeleman and Titov 2009, Bryzgunova 1980, Yokoyama 1986, Krylova and Khavronina 1988, Meyer and Mleinek 2006). Discourse-new referents were characterized as prosodically augmented through greater intensity, pitch range and duration, compared to discourse-given or highly inferable referents (Luchkina and Cole under review).

Sekerina and Trueswell (2012) argue that prosody in Russian acts independently of the word order (for example, under canonical SVO configuration), or 'in consort with word order variation', in spoken discourse. When the canonical SVO order is preserved, location of the nuclear pitch accent in the sentence may vary if a non-sentence-final constituent is under focus. Jasinskaja (2013)

[15] Neeleman and Titov (2009, footnote 2) characterize contrastive foci in Russian are higher in tone and prosodically augmented. Sekerina and Trueswell (2012) report that contrastive foci in Russian are obligatorily pitch-accented, as evident from the lengthening of the sressed syllable of the focused word in combination with the low boundary tone (L%).

concludes that In Russian, just like in English and other prominence dislocating languages, monitoring the location of the nuclear pitch accent is highly important for determining the location and size of the focused constituent.

The interaction of intonation and word order characteristics in Russian makes it fit the definition of a prominence-dislocating and a constituent-dislocating language. Further complicating this picture, Sekerina and Trueswell (2012), Botinis et al. (2005) and Luchkina and Cole (2016) demonstrated that a change in word order in Russian is coincidental with a change in the acoustic-prosodic expression of the ex-situ word. This change is manifested though systematic variation in acoustic-prosodic expression which results from two kinds of constituent dislocation to clause-peripheral positions: fronting a word to the clause-initial position and post-posing a word to the clause-final position.

To illustrate, in a study of focus correlates in Russian and Greek, Botinis, Kostopoulus, Nikolaenkova and Themistocleous (2005) used 480 elicited Greek utterances and 720 Russian elicited utterances to determine the relative contribution of syntactic and tonal (acoustic-prosodic) cues in the expression of broad and narrow focus in these languages. Botinis et al. (2005) documented numerous instances of 'syntactic dislocation' of the focused word to the rightmost sentence boundary in Russian (and the leftmost sentence boundary in Greek, pp.100-101). They documented that robust tonal changes, specifically, a local tonal range expansion followed by a tonal compression of the post-focal material, marked the narrowly focused in- and ex-situ constituents in Russian. They tentatively proposed that prosodic and syntactic correlates of focus in Russian complement each other and are mutually reinforcing. Luchkina and Cole (2016) compared acoustic-prosodic expression of ex-situ fronted and post-posed words to that of their in-situ counterparts using read production data from fifteen Russian speakers. They reported that in

comparison to a control sample of in-situ words, ex-situ fronted words demonstrated systematic increases in pitch range, segment mean intensity and duration. Ex-situ post-posed words, on the contrary, had reduced pitch range but greater segment duration. Evidence presented by Botinis et al. (2005) and Luchkina and Cole (2016) points to the possibility that structural prominence and prosodic prominence are used simultaneously in Russian and both contribute to the likelihood that a word is perceptually prominent in discourse.

The present study

Although the roles of prosodic accenting and constituent linearization in discourse have been considered, separately, in the previous studies, the effects of these mechanisms jointly available during discourse comprehension have not yet been experimentally tested. No study known to us has addressed the perceptual significance of word order and acoustic-prosodic cues for discourse comprehension in languages where constituent *and* prominence dislocation mechanisms are available and may be engaged simultaneously. In the absence of prior experimental work, the status of the concurrent application of constituent linearization and prosodic cues in relation to perceived prominence remains largely unclear. Listeners may choose to attend to one of these cues based on their relative validity in a given context, or use constituent ordering and acoustic-prosodic information simultaneously.

The present study aims to start filling the gap in understanding how word order and prosody work separately or together to express prominence in Russian, a language, where word order and prosodic variability may be simultaneously available during discourse comprehension. We test if variation in word order and acoustic-prosodic cues serves as a means of expressing relative accessibility and information status of a word and, by doing so, mediates its perceived prominence.

In what follows, we examine variability in word order and prosodic parameters which increases perceived prominence of a word in discourse and tap into the relationship between these two distinct prominence encoding mechanisms. We begin by gathering confirmatory evidence for two distinct sources of acoustic-prosodic variation in read discourse in Russian. One of these sources is related to the information status of a word, which we use as a proxy to that word's inherent and perceived prominence. The second source of acoustic-prosodic variation is related to a non-canonical linearization of sentence constituents. We focus on two kinds of non-canonical or ex-situ positions, clause-initial and clause-final, reflective of constituent fronting and post-posing. Both these types of surface constituent movement have been attested for Russian and are of special interest since they align the ex-situ constituent with a prosodic domain (intonation phrase) boundary. Domain boundary positions are known to be inherently perceptually salient; proximity to such positions, cross-linguistically, correlates with prominence and may serve as a motivation for constituent displacement (Féry 2013). Consistent with this view, Gow, Melvold, Manuel (1996) and Cho (2016) characterize domain-initial positions as informationally-rich and articulatory strengthened, cross-linguistically. We believe the same assumptions can be plausibly tested for domain-final (also clause-final) positions in Russian, since such positions are nuclear pitch accented and often associated with discourse-novel information. In line with the previous work by Luchkina and Cole (2016) and Botinis et al. (2005), we seek confirmatory evidence for prosodic strengthening of ex-situ clause-initial and clause-final constituents and gauge if ex-situ elements, in general, are perceived as highly prominent. To this end, we ask linguistically naïve native Russian speakers to rate word prominence in read and spoken discourse. We then evaluate the relative cue validity of prosodic and word order-related prominence marking mechanisms in

Russian by comparing their functional significance during auditory comprehension of read discourse.

Our results support the following important observations. Two independent sources of prosodic variability are observed in read discourse. One of these sources pertains to an ex-situ position in a sentence or phrase and its proximity to an intonation phrase boundary: ex-situ fronted words are prosodically augmented, whereas ex-situ post-posed words, although lengthened, appear prosodically reduced. The second source of prosodic variability is accent-induced and pertains to prominence maximization and greater perceptual saliency of in-situ words. This second kind of prosodic variability reflects the information status and relative accessibility of words and their referents in discourse. More specifically, we find that prosodic augmentation marks discourse-novel information.

Prosodic augmentation is positively associated with the probability that a word is identified as prominent by the listener. Similarly, ex-situ words, regardless of the direction of the movement which originated them, are likely to be perceived as prominent. These findings point to the possibility that the relationship between constituent linearization and prosodic augmentation in Russian is that of mutual reinforcement. This conclusion, however, needs to be qualified due to a special acoustic-prosodic status of ex-situ post-posed words. Despite being nuclear pitch-accented and bearing an association with discourse-new information, post-posed words feature reduced intensity and compressed pitch range. We argue that these acoustic-prosodic changes may be production-specific or pertain to prosodic landmarks, such as domain boundaries. Listeners seem to be aware of the nature of acoustic-prosodic variability concurrent with a change in word order and use prosodic expression of ex-situ words probabilistically. This conclusion is based on the finding that prosodic augmentation of ex-situ fronted words is positively correlated with the

likelihood that these words are rated as prominent. At the same time, prosodic reduction of ex-situ post-posed words is unrelated to their perceived prominence status, possibly due to the inherent perceptual salience of the sentence-final nuclear pitch-accented position that cannot be offset by articulatory weakening of the phrase-final word (Ayers 1996, Terken and Hermes 2000). Comparison of the functional significance of word order and prosodic cues in Russian reveals that the latter explain relatively more variability in the prominence ratings obtained during read discourse comprehension. Acoustic-prosodic and word order factors, combined, yield the best modeling of perceived prominence.

Production data

Because a word, spoken or written, is usually normally perceived as prominent in context, we begin our investigation with a series of descriptive analyses of word order and acoustic-prosodic variability in two discourse samples chosen for this study. These materials come from two published narratives in Russian which we analyzed in our previous work (Luchkina and Cole (2016). Luchkina and Cole examined read productions of these narratives by 15 native Russian speakers and documented acoustic-prosodic characteristics of nominal expressions related, among other things, to a (non-canonical) position of a word in the sentence and its referent information status. Our earlier analyses confirm that these narratives are highly representative of word order and acoustic-prosodic variability in discourse, which in this study we anticipate to contribute to perceived prominence in Russian.

Our first discourse sample (Text 1) presents an excerpt from a biography of a Russian poet; the second discourse sample (Text 2) is an unabridged folk tale. Two stylistically different texts were chosen to reflect more standard (Text 1) and more colloquial (Text 2) language use. Cumulatively, these materials contain 344 content and 69 function words. The average sentence length in both

narratives is 5.2 content words (SD=1.77); 20% of all clauses and 18% of all content words occur ex-situ, in non-SVO clauses. Text 1 includes 29 SVO clauses, 3 OVS, and 1 SOV clauses. Text 2 includes 25 SVO, 3 OSV, 2 VSO, 2 SOV, 4 OVS, 2 OV, and 3 VS clauses. Such uneven distribution of word orders in the narratives is not unusual: In Russian, SVO is the basic, pragmatically neutral word order compatible with all information structure configurations; OVS presents the more common non-canonical word order. Following Luchkina and Cole (2015), each content word position was coded as *in-situ*, *fronted* or *post-posed*, relative to the canonical SVO order, for a total of 28 fronted and 33 post-posed words.

Information status distinctions as a prominence scale

Per Jasinskaja (2013), one function that word order variability performs in Russian, as well as cross-linguistically, is that of an information structure preserving mechanism. This mechanism operates via optimizing the distribution of information of different degrees of accessibility or givenness in a sentence or phrase, such as to optimize production and processing (Arnold et al. 2000). Acoustic-prosodic variability is another widely known correlate of information structure, particularly, in languages with limited word order freedom (Breen et al. 2010).

Numerous studies, including Arnold et al. (2000), Baumann and Riester (2012), Baumann (2006), Breen et al. (2010), Skopeteas and Fanselow (2008), associate highly accessible given information with prosodic reduction and frequent pronominalization, indicative of its lower information impact in discourse. Discourse-novel information, on the contrary, is typically prosodically augmented and resists prominalization. Based on this evidence, we anticipate that observable variation in prosodic expression and constituent ordering may reflect the information status and relative accessibility of a lexical word and/or its referent. For this reason, we deem information structure a

relevant prominence scale for this study and treat information status of a word and its referent as an anchor marker of that word's perceived prominence.

Some widely used information structure frameworks operate three-way distinctions of information (e.g., Chafe (1976), Prince (1981), Calhoun, Nissim, Steedman and Brenier (2005)). Typically, these frameworks call for distinguishing between information which is *discourse-novel* (1st mention in the narrative, least accessible and most prominent information), *discourse-given* (repeated mentions in the narrative, most accessible and least prominent information), and *discourse-inferable* (1st mention in the narrative in combination with being highly inferable and less prominent). In this study, we determine the information status of a word at two distinct levels, referential and lexical, specified under RefLex annotation scheme for information structure proposed by Baumann and Riester (2012, 2013) and summarized below.

Rooted in Chafe's cognitive approach to information structure (1976, 1994), RefLex offers a rather detailed discrimination of information categories at two distinct levels, *referential*, pertaining to the properties of the word referent and *lexical*, pertaining to the lexical choices a speaker makes to identify a referent in discourse[16]. Following Bauman and Riester (2012), at the referential level of

[16] RefLex allows to successfully capture instances in discourse where co-referential expressions are used to refer to the same discourse entity (*Tatiana's father* mentioned in (3.7a) below is referred to as *Esenin's grandfather* in example (3.7b) or where the same lexical item is used to identify a novel discourse referent (as shown in (3.8a) and (3.8b), within a narrative, the word 'father' is used to identify two different referents, *Tatiana's father* and *Sergey's father*).

(3.7a) Tatiana ushla k svoemu **otcu**
Tatiana-NOM went to her-DAT father-DAT
Tatiana left (her husband) to stay at her father's house.

(3.7b) **Ded** **Esenina** byl yarkoj lichnost'yu. grandfather
Esenin-GEN was bright personality Esenin's
grandfather had an interesting personality.

annotation (*r-*), we classified referents in the narratives as *r-given*, *r-bridging*, *r-new*, and *r-unused*. At the lexical level of annotation, we focused on the lexical inventory in the narratives and classified content words as *l-given-same*, *l-given-synonymous*, *l-accessible*, or *l-new*. These referential and lexical information categories are defined in Table 3.1. Assignment of select IS categories, at lexical and referential levels of annotation is illustrated in (3.9).

Table 3.1. *Definitions of information categories from RefLex scheme by Bauman and Riester (2012, 2013): Top panel: referential information categories; bottom panel: lexical information categories.*

Referential level	*r-given:* anaphor corefers with antecedent in previous discourse	*r-bridging:* non-coreferring anaphor, dependent on preceding context	*r-new:* a new referent/concept	*r-unused-known* (henceforth: r-unused) discourse-new item which is generally known (e.g, a toponym)
Lexical level	*l-given-same:* recurrence of same expression	*l-given-synonymous:* synonyms	*l-accessible:* two lexically-related words, (e.g., hyponym-hyperonym)	*l-new:* word unrelated to another word within last 5 intonation phrases or clauses

(3.8a) Tatiana ushla k svoemu **otcu**
Tatiana-NOM went to her-DAT father-DAT Tatiana left (her husband) to stay at her father's house.

(3.8b) **Otec** Sergeya, Aleksandr, sluzhil v Moskve.
father-NOM Sergey-GEN Alexander-NOM served in Moscow
Sergey's father, Alexander, held an office appointment in Moscow. (Examples from Text1).

(3.9) Preceding context:

Once a fox ran along a road…and fell into a well. There was not much water in the well. The fox did not drown but could not get out, either. The fox was sitting in the well and thinking of what could be done to remedy the situation. At this time, …

…po	doroge	shol	kozyol	
along	road	walked-M	goat-NOM	
r-given		r-bridging	r-new	*Referential IS label*
l-given-same		l-accessible	l-new	*Lexical IS label*

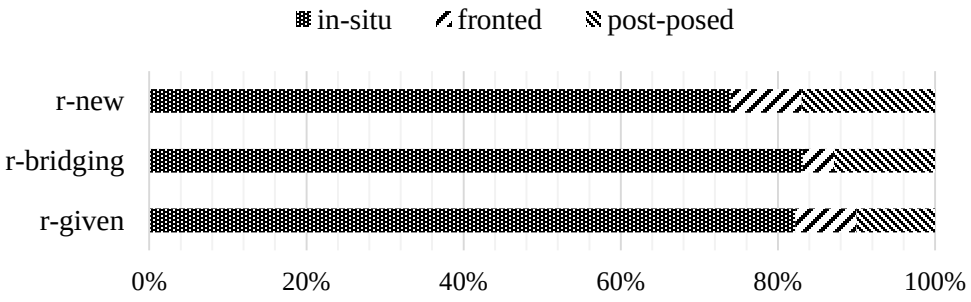
'…a goat walked along the road.'

In our earlier work (Luchkina and Cole 2016), we determined that the discourse samples used in the present study demonstrate a meaningful distribution of discourse referents of different information status across the three types of sentence positions which are of interest to this study, in-situ, ex-situ fronted, and ex-situ post-posed. In our earlier analysis, discourse-novel referents proved the most mobile referent class in the narratives and occurred in all sentence positions and accounted for most instances of ex-situ post-posed words. Discourse-given and discourse-inferable referents, on the contrary, were the more positionally-grounded referent categories occurring mostly in-situ or fronted, relative to the canonical position, as shown in Figure 3.1. This distribution agrees with the assumption that sentence-initial position in Russian bears an association with discourse-given information (Neeleman & Titov 2009, Slioussar 2011b, Ionin & Luchkina, under review), whereas the sentence-final, nuclear pitch-accented position is associated with discourse-novel information. Such distribution of referent categories also complies with the cross-linguistically attested information structure template, whereby discourse-given information and information accessible to the speaker and the listener, precede discourse-novel information.

For the purposes of this study, information status of each content word was evaluated at separately at the lexical and referential levels and was represented by two information status variables,

IS_REF, and IS_LEX. In the variable names, 'IS' stands for 'information status' and '_REF' and '_LEX' mark the relevant annotation level (see Table 3.1.). The annotation was performed by one of the authors (TL) and another native Russian speaker, for a total of 344 content words[17]. Inter-rater agreement (linearly weighted Kappa) between the annotators and across Text 1 and Text 2 was satisfactory: $\varkappa$=0.89, *SE*=0.03, α=0.05. Two levels of IS_LEX, *l-given-same* and *l-given-synonymous* were found to be collinear during statistical modeling. The collinearity issue was resolved by merging words annotated as *l-given-same* and *l-given-synonymous* under a joint category *l-given*.

Figure 3.1. *Distribution (%) of referential information categories across sentence positions in PRT materials. Y-axis: Referent information status. X-axis: % occurrences.*



Read production data analyses and acoustic features pre-processing

In this section, we focus on the reading performance of a female native speaker of Russian whose read production was used in the unguided prominence rating task reported below. The model speaker, age 27, acquired Russian from birth and resided in Russia until age 26. At the time of participation she was a graduate student in the Slavic Department of a public University located in

[17] Baumann and Riester's framework extends to non-nominal expressions, including verbs and adjectives. Since in the prominent rating task, which presents the key experimental manipulation reported in this work, raters were not restricted in their prominence ratings to any specific grammatical category of words in discourse, we submit data for *all* content words in the materials to analyses reported in this and subsequent sections.

U.S. Midwest. The model speaker was chosen based on her ability to read expressively and with satisfactory clarity. She was instructed to read the narratives in a lively naturalistic manner, as if addressing an audience. All recordings were made in a sound-proof recording booth using a Marantz PDM 750 solid state recorder and a head-mounted microphone. Recorded data were digitized at the sampling rate of 44 kHz and subject to acoustic analyses. The acoustic measures of f0 maxima and range (indicative of f0 excursion size, per Traunmüller and Eriksson 1995), mean intensity and duration were taken from every syllable[18] of each content word in the narratives.

All measurements were extracted automatically in Praat (Boersma and Weenink 2016). Fundamental frequency and intensity measurements were taken from the center region of the vowel in order minimize the influence of the adjacent segments at the voice onsets and during inter-segmental transitions. Each f0 output was transformed to a semitone value relative to a fixed value of 100 Hz. For the purposes of statistical analyses, log transformation and mean-centered coding of acoustic-prosodic measures was implemented (Mitchell 2012, Bush, Hess, and Wolford 1993). Extracted acoustic measures, summarized in Table 3.2., were then examined as correlates of information status, and analyzed for their relationship to position of a word in the sentence.

Model speaker's reading performance was carefully examined for evidence of the interaction between information structure, word order, and prosodic variability in the study materials. We focused, specifically, on the phonetic manifestations of (a) information structure distinctions and (b) ex-situ position of a word in a sentence or phrase.

[18] A syllable, rather than a word, was chosen as the appropriate unit for extracting acoustic-prosodic measurements to account for prosodic events occurring in the post-stressed vowel region (aka post-tonic pitch movements in Russian (Makarova 2007, Svetozarova 1998).

Table 3.2. *Summary statistics of acoustic-prosodic parameters extracted from the model speaker's read production data (mean, SD, range).*

Word Order	**fronted** (n=28)	min-max	**in-situ** (n=283)	min-max	**post-posed** (n=33)	min-max
duration (ms)	92.31 (32.4)	142.7	79.99 (46.2)	357.9	91.04 (43.5)	193.7
mean intensity (dB)	75.86 (4.7)	22.8	75.16 (4.9)	46.9	74.2 (5.1)	20.1
pitch range (Hz)	33.16 (7.88)	28.3	31.2 (7.6)	42.4	29.0 (10.3)	39.16
max f0 (Hz)	296.9 (82.3)	320.1	274.0 (71.3)	377.1	258.8 (67.3)	330.3

IS REF	**r-bridging** (n=55)	min-max	**r-given** (n=96)	min-max	**r-new** (n=193)	min-max
duration (ms)	68.15 (26.9)	146.2	85.4 (46.9)	357.9	84.9 (48.2)	328.6
mean intensity (dB)	75.1 (4.36)	27.8	74.44 (5.5)	44.8	76.5 (4.7)	34.7
pitch range (Hz)	35.2 (34.9)	30.9	41.4 (38.6)	36.1	42.3 (45.8)	40.47
max f0 (Hz)	271.6 (65.8)	217.0	273.9 (76.2)	362.4	275.5 (72.4)	444.2

IS LEX	**l-accessible** (n=97)	min-max	**l-given, given-synonymous** (n=81)	min-max	**l-new** (n=166)	min-max
lduration (ms)	74.75 (33.8)	180.8	87.50 (49.8)	357.9	83.64 (48.0)	328.7
mean intensity (dB)	74.59 (5.51)	46.8	74.66 (4.5)	25.5	75.64 (4.66)	34.7
f0 range (Hz)	30.73 (8.9)	41.7	31.14 (8.3)	34.5	31.37 (7.3)	44.7
max f0 (Hz)	270.5 (76.1)	387.4	276.6 (75.9)	362.4	275.5 (68.3)	428.6

Confirmatory hypotheses and predictions

Prosody, through systematic variation in acoustic parameters, has been characterized as a cue to information status and relative information accessibility in Russian (Svetozavora 1998, Jasinskaja 2013). More recently, Botinis et al. (2005) and Luchkina and Cole (2016) and documented another case of prosodic variability in the language, pertaining to the non-canonical linearization of sentence constituents. In Botinis et al.'s analysis of semi-elicited data, focused ex-situ words we characterized by a local tonal range expansion in combination with compression of the post-focal material. Luchkina and Cole (2016) reported acoustic augmentation of ex-situ fronted words in combination with partial acoustic reduction of ex-situ post-posed words in the read production of

fifteen Russian speakers. These acoustic-prosodic changes were observed regardless of the referent information status. Based on the record of acoustic-prosodic variability observed in conjunction with information status and constituent ordering in the previous work, we seek confirmatory evidence for the following hypotheses.

Hypothesis 1: *Acoustic-prosodic variability in the model speaker's reading performance reflects information status of a word, at lexical and referential levels.*

Hypothesis 2: *Acoustic-prosodic variability in the model speaker's reading performance reflects constituent ordering in a sentence or phrase.*

Understanding if the variation in the acoustic-prosodic parameters in the model speaker's reading performance reflects information structure dynamics and/or constituent ordering will provide important insights into the prosodic information available to the listeners during discourse comprehension and will help us interpret results of the auditory prominence rating task reported below.

Statistical modeling of the read production data

To test confirmatory hypotheses (1) and (2), model speaker's read production data were examined for acoustic-prosodic variation indicative of (a) the information status of a word relative to other words in a sentence or phrase and (b) the position which a word takes relative to the canonical SVO order (in- vs. ex-situ). To capture the simultaneity and interrelatedness of prosodic processes during speech production, we built a system of simultaneous multivariate linear regressions, which we will refer to as Model 1.

Model 1 includes four dependent variables, each one corresponding to an acoustic-prosodic parameter of interest *mean intensity*, *duration*, *f0 range,* and *max f0* and the independent variables *Word Order*, *IS Lex* and *IS Ref.* All variables and their levels are summarized in Table 3.3.

Table 3.3. *Independent variables examined in the analysis of the model speaker's read production. The reference level for each variable is highlighted in bold.*

Independent variables	**variable levels**
Word Order	**in-situ**
	ex-situ, fronted
	ex-situ, post-posed
Information status, referential level (IS_REF)	**r-given**
	r-bridging
	r-new
	r-unused
Information status, lexical level (IS_LEX)	**l-given**
	l-new
	l-accessible

The output of Model 1 is shown in Table 3.4. Each panel of Table 3.4. summarizes the effects of predictor variables on one acoustic-prosodic parameter of interest, that is, one of the four dependent variables.

Results

Prosodic effects of information status

Let us first examine how prosodic parameters fundamental frequency, intensity, and duration co-vary with the information status of a lexical word and its referent, when word order is controlled for. Results of the multivariate regression analysis reveal that relative to the IS_REF category *r-given*, words which introduce discourse-novel referents (*r-new*) are prosodically augmented. Notably, all examined prosodic parameters undergo augmentation when a discourse referent is new, mean intensity (t=5.75, p<.001), duration (t=2.36, p<.05), f0 range (t=3.58, p<.001), max f0 (t=2.74, t<.01) This finding agrees with the cross-linguistically established fact that discourse-new information status is indicative of greater perceived prominence.

Table 3.4. *Results of the multivariate regression analysis of the model speaker's production data.*

Fixed effects	**mean intensity**	β	SE	P>| *t* |
Word Order	fronted	.12	.03	<.001
	post-posed	-.17	.03	<.001
IS_REF	r-unused	-.72	.2	<.001
	r-bridging	-.24	.03	<.001
	r-new	.21	.04	<.001
IS_LEX	l-accessible	-.17	.03	<.001
	l-new	-.01	.04	n.s.
	duration			
Word Order	fronted	.22	.03	<.001
	post-posed	.23	.03	<.001
IS_REF	r-unused	-.76	.19	<.001
	r-bridging	-.23	.03	<.001
	r-new	.08	.04	<.05
IS_LEX	l-accessible	-.21	.03	<.001
	f0 range			
Word Order	fronted	.23	.03	<.001
	post-posed	-.01	.02	n.s.
IS_REF	r-unused	.327	.2	n.s.
	r-bridging	-.04	.03	n.s.
	r-new	.13	.03	<.001
IS_LEX	l-accessible	-.16	.03	<.001
	l-new	-.17	.03	<.001
	max f0			
Word Order	fronted	.29	.03	<.001
	post-posed	-.24	.29	<.001
IS_REF	r-unused	1	.19	<.001
	r-bridging	.09	.03	<.01
	r-new	.07	.03	<.01
IS_LEX	l-accessible	-.13	.03	<.001
	l-new	-.06	.04	n.s.

Prosodic effects associated with the intermediate referent categories *r-unused* and *r-bridging* imperfectly pattern with *r-new* information: both these information statuses are associated with greater f0 maxima and mean intensity than *r-given* (*max f0*: t=5.15, p<.001 for *r-accessible* and t=7.25, p<.01 for *r-bridging*; *mean intensity*: t=3.71, p<.001 for *r-accessible* and t=2.65, p<.01 for *r-bridging*) and smaller duration than *r-given* (*r-accessible*: t=3.91, p<.001; *r-bridging*: t=-7.11, p<.001).

These results are consistent with Baumann (2006), who reported that intermediate information categories in German were prosodically accented, however in a more subtle way (L* accent, using ToBi notation), indicative of their semi-active status in discourse.

Turning now to the analysis of the lexical information categories represented by the variable IS_LEX, results of the multinomial regression reveal that *l-accessible* and *l-new* words are overall more prosodically reduced than the *l-given* words. In more detail, *l-accessible* words appear consistently prosodically reduced, by virtue of having lower mean intensity (t=-5.36, p<.001), reduced duration (t=-6.66, p<.001), smaller f0 range (t=-5, p<.001), and lower f0 maxima (t=-4.19, p<.001) than *r-given* words. Prosodic reduction of *l-new* words is partial, as evident from smaller duration (t=-4.59, p<.001), and smaller f0 range (t=-4.47, p<.001) in comparison with *l-new* words (see Table 3.3.). These results reveal that *l-new* and *l-given* words in the model speaker's read production cannot be discriminated based on mean intensity or f0 maxima, and that *l-new* words' partial acoustic-prosodic reduction may reflect their referential status. The departure from the baseline category *l-given* is greatest for the *l-accessible* words in our materials. The latter feature significantly lower pitch range and maxima, mean intensity, and duration than the *l-given* words. We conclude that prosodic characteristics of the lexical information categories are far less straightforward than those observed for the referential information categories. These results

provide partial support for Hypothesis 1, which posits that acoustic-prosodic variability in the model speaker's reading performance cues information status and relative accessibility of a (lexical) word and that word's referent.

Prosodic effects of word order

Analysis of the model speaker's read production data revealed that a number of prosodic changes systematically distinguish ex-situ words in the model speaker's read production data. To illustrate, words fronted to the sentence-initial position appeared prosodically augmented and featured significantly greater mean intensity (t=4.09, p<.001), duration (t=7.36, p<.001), and pitch range (t=7.54, p<.001) and maxima (t=9.49, p<.001). Words post-posed to the sentence-final position, on the contrary, were partially reduced, as evident from lower mean intensity (t=-5.8, p<.001) and f0 maxima (t=-8.61, p<.001). At the same time, post-posed ex-situ words had significantly greater segment duration than their in-situ counterparts (t=8.04, p<.001). These findings are consistent with word order-specific patterns of acoustic-prosodic variation we detected in the read production of fifteen Russian speakers in our earlier work (Luchkina and Cole 2016). These results also support Hypothesis 2, which predicts that acoustic-prosodic variability in the model speaker's reading performance is reflective of non-canonical constituent ordering in discourse.

Interim summary

So far, we have examined acoustic-prosodic variation in the model speaker's read production in relation to information status of a word and its position in a sentence or phrase. Prosodic variability related to these factors is of interest, as it reveals what prosodic information is available to the listener during auditory comprehension.

In a system of simultaneous linear equations, we examined the psycho-acoustic parameters length, loudness and pitch, represented by acoustic-prosodic measures duration, intensity, and fundamental frequency extracted from the model speaker's read production data. We examined variation in these acoustic-prosodic parameters in relation to the information status of a lexical word and its referent. We then examined the variation specific to ex-situ words, namely words which are fronted or post-posed, relative to the canonical SVO order.

Our analyses yielded confirmatory evidence for Hypothesis 1, which was motivated by the patterns of acoustic-prosodic variation in the read performance of fifteen native Russian speakers reported in Luchkina and Cole (2016). Hypothesis 1 predicted variation in the acoustic-prosodic expression specific to the word's information status and relative accessibility in discourse. We found that information structure dynamics, whereby words of different lexical and referential status get (re-)introduced into discourse, brings about a number of prosodic changes. To illustrate, in the model speaker's reading performance, referentially-new information was prosodically augmented, consistent with rich cross-linguistic evidence that discourse-new information is prosodically and perceptually prominent in discourse. Easily inferable and therefore more accessible *r-unused* and *r-bridging* information showed evidence of partial prosodic reduction by virtue of having shorter duration. This result is expected, given the lower information load and higher overall predictability of such information in discourse. *R-unused* and *r-bridging* referents also shared a number of prosodic features with discourse-new referents; specifically, their mean intensity and pitch maxima were greater than those of the *r-given* referents. These findings point to an intermediate prosodic status of relatively more accessible unused and bridging referents, due to a counteracting influence of their greater predictability in combination with the fact that these referents are introduced into discourse for the first time (and, technically, represent a sub-category of *r-new* information).

The lexical level of RefLex focuses on the information status of lexical words rather than their referents. Prosodic variation which we observed in conjunction with the lexical information categories was less straightforward. For example, contrary to the expectation expressed by Baumann (2006) that discourse-new information should be accented and therefore appear prosodically prominent (relative to discourse-given information), *l-given* words in the read production of our model speaker were, in a number of aspects, more prosodically strengthened than *l-novel* words. A more expected result was obtained for *l-accessible* words, which, by virtue of having high cloze probability, were de-accented.

We conclude that the observed acoustic-prosodic variability unequivocally supports the special status of discourse-novel referents whose prosodic prominence was supported by multiple acoustic parameters in the model speaker's read production. Our results, overall, warrant the differentiation of two distinct levels of the information status, lexical and referential, per Bauman and Riester's proposal (2012, 2013). However, since in this work we only examine production of one speaker, prosodic cues which support this division in Russian merit further investigation.

A qualitatively distinct source of acoustic-prosodic variability in the model speaker's production reflects whether a word in the narratives occurs in- or ex-situ in the sentence or phrase. Fronting a word relative to its canonical position boosts a number of acoustic-prosodic parameters associated with that word, including mean intensity, duration, and pitch range and maxima. Prosodic effects indicative of post-posing a word to the sentence-final position include augmenting its duration, indicative of a more pronounced phrase-final lengthening, in combination with lowered mean intensity and pitch maxima, and compression of the pitch range. In summary, we found that a change in word order triggers a series of prosodic effects which largely depend on whether the ex-

situ word is fronted or post-posed relative to its canonical position. These results present confirmatory evidence for Hypothesis 2.

Analyses of the read production data helped us establish that acoustic-prosodic parameters intensity, duration, and fundamental frequency extracted from the model speaker's read production data vary in relation to the information status of a word and are *concurrently* affected by that word's position in a sentence or phrase. These results are consistent with the proposal that word order and prosodic cues to perceived prominence may be simultaneously available in a free word order language and may jointly contribute to a word's perceived prominence. Next, we put this proposal to an empirical test.

Unguided prominence rating task (PRT)

Analyses of the read production data demonstrated that argument linearization and information status of a word jointly affect its prosodic expression. We documented that both these factors may trigger acoustic-prosodic augmentation and reduction processes. Consistent with these findings, in this section we test three discourse components which may contribute to perceived prominence in Russian. These components are

- Information structure (with focus on information status of a word as an anchor weight for its relative prominence in discourse);

- Linear order of words in a sentence or phrase relative to the canonical SVO order (with focus on two kinds of ex-situ positions, ex-situ fronted and ex-situ post-posed);

- Acoustic-prosodic expression (with focus on observable co-variation in acoustic parameters with information status of a word and its position in a sentence or phrase).

Method

In an unguided prominence rating task (henceforth, PRT) Russian speakers (n=81) assessed the relative prominence of content words while reading or listening to the discourse samples examined in the previous section. Two PRT versions were administered, a silent reading PRT and an auditory PRT. The key difference between these versions was modality of the discourse, written, with no acoustic-prosodic cues vs. auditory, in which raters listened to the recording of the model speaker read production while performing the rating task. Consistent with this design, we anticipate that information status should serve as a stable correlate of perceived prominence, regardless of the PRT modality. We also anticipate that word order and prosody may contribute to perceived prominence ratings differently in the two PRT versions. We treat the silent reading PRT results as a baseline estimation for the functional significance of word order, which we refer to as a structural cue to perceived prominence in Russian. Results of the auditory PRT are more complex, in that they reflect the relative contribution of the model speaker's prosody to perceived prominence. Auditory PRT responses allow us to gauge the relative contribution of word order and prosodic cues to perceived prominence and compare the functional significance of these prominence encoding mechanisms in Russian.

Each PRT version included thirty-nine clause-size excerpts from Text 1 and Text 2. Depending on the task modality, each excerpt was presented as written text or an audio recording of the model speaker's reading performance. A clause was chosen as a unit of presentation because it expresses one relatively complete idea and can be perceived as a whole. Each discourse segment was presented along with the preceding context provided in the written form at the top of the experiment page. Participants read the entire portion of the text preceding the target segment, read (in the silent reading PRT) or listened to (in the auditory PRT) the target segment and identified

the prominent word(s) in each segment by associating them with one level of the binary feature "+/- prominent". In the silent reading modality, participants placed a tick mark next to each word which they perceived as prominent. In the auditory modality, participants listened to the target segment as many times as they wished to and entered their responses into a textbox. For each audio clip, they were provided with unlabeled (numbered) text boxes equivalent in number to the number of the content words in the target segment. Following Cole, Mo and Hasegawa-Johnson (2011), no formal definition of prominence was given. Participants were instructed to identify only those words as prominent that 'were the focus of their attention'. Participants could choose as few as one word or as many as all of the content words in the target segment as prominent.

Participants

Forty-nine native Russian speakers (ages 18-42) completed the silent reading PRT and 32 speakers (ages 19-38) completed the auditory PRT. At the time of participation, all participants resided in Russia and reported Russian as their native language, as well as the only language spoken in their household. Participants accessed the task online from a custom-built test page hosted by the University of Illinois server.

Hypothesis and Predictions

If Russian patterns like languages in which only one cue to perceived prominence, prosody or word order, is used to the exclusion of the other cue (per Swerts et al. 2002 and Donati and Nespor 2003), we expect prominence ratings obtained in the PRT experiment to be affected by acoustic-prosodic *or* linearization factors only. If, on the contrary, Russian deploys prosodic variability and relative constituent ordering as cues to perceived prominence, we expect raters to show sensitivity to word order *and* acoustic-prosodic variability alike.

In the previous section, we reported that prosody and word order are jointly involved in the expression of information structure relations in Russian and that ex-situ words have tend to be prosodically distinct. These findings lead us to formulate the following research hypothesis:

Hypothesis 3: *An ex-situ sentence position acts as an independent cue to prominence in Russian and is further reinforced with acoustic-prosodic features associated with such position.*

Hypothesis 3 generates three straightforward predictions notated as (b), (c), and (d), regarding the role of acoustic-prosodic cues and word order cues in Russian, listed below. Additionally, we would like to formulate a default prediction (a) as a validation test of the PRT methodology.

(a) We expect to observe a meaningful dissociation in the prominence ratings obtained for different information categories represented in the PRT materials. Specifically, we expect that discourse-given words have a lower probability of being rated as prominent, whereas discourse-new words have a greater probability of being rated as prominent, per Breen et al. (2010), Jasinskaja (2013) and Baumann and Riester (2012, 2013). We anticipate this pattern of results to hold regardless of the PRT modality. By testing this default prediction, we will be able to confirm that words chosen as prominent in this study share a known prominence-related criterion established in the previous research, information status.

Moving on to modality-specific predictions, *in the Silent Reading PRT*, per our hypothesis, we predict (b) that ex-situ words will have a greater likelihood of being rated as prominent, since during the silent reading, word order variability is a readily available cue to prominence, whereas prosodic expression has to be supplied by the reader in the form of 'silent prosody'. *In the Auditory PRT*, word order variability is supplemented with the model speaker-generated prosody expression. Per Hypothesis 3, we therefore predict that (c) raters will identify words which are

prosodically augmented as prominent and (d) that raters will identify words which occur ex-situ as prominent.

Additionally, the PRT design allows us to evaluate the relative functional weight of constituent linearization and prosodic cues to perceived prominence by comparing the relative contribution of each cue type to obtained prominence ratings.

Statistical modeling of the PRT responses

The PRT responses from each rater were coded as a binary variable, *p(rominence)-score*, which equaled 1 if a rater selected a word as prominent, and 0 otherwise. The inter-rater agreement coefficients were computed to assess the level of agreement among raters, in each modality. Obtained Fleiss' kappas translate into fair but highly significant agreement levels: silent reading PRT κ=0.26, p<.001 and auditory PRT κ=0.36, p<.001. We conclude that although respondents were not given a definition of prominence, their performance during both PRT experiments was above chance.

We will now examine the prominence ratings, or p-scores, obtained in the two PRT versions, based on the modality in which the discourse samples were available to the raters. While the role of information structure is evaluated in both PRT versions, the modeling of prominence scores obtained in the modality-specific PRTs modality-specific. In the silent reading PRT, we examine if position of a word in a sentence or phrase contributes to its perceived prominence. In the auditory PRT, where both word order and acoustic-prosodic cues were simultaneously available to raters, we test the relative contribution of word order and acoustic-prosodic cues to perceived prominence in Russian. In line with predictions (a-d), we first summarize and visually illustrate the p-scores obtained in the auditory and silent reading PRTs. For ease of exposition, following Mo, Cole, and

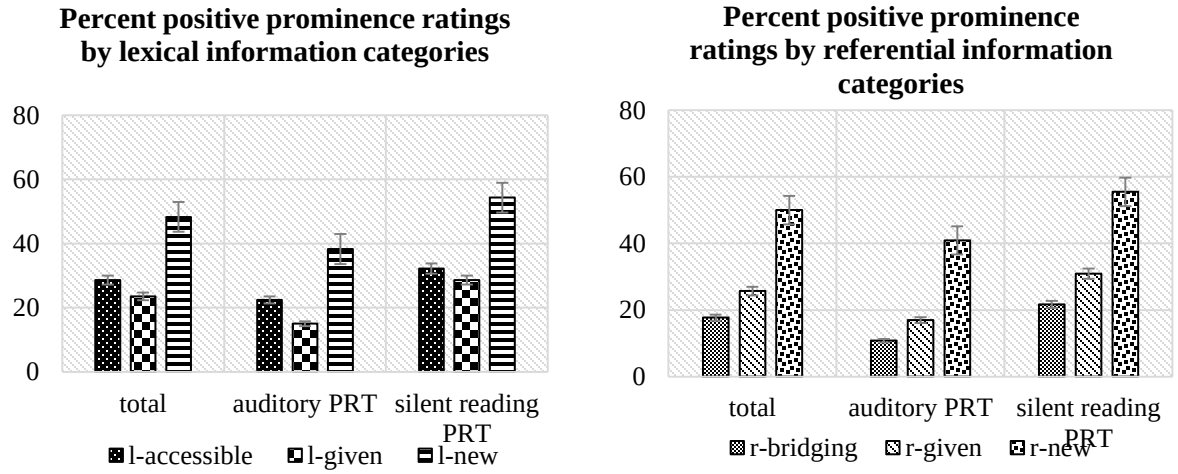
Lee, (2008), for each content word in the narratives, we computed percent of respondents who marked it as prominent, per PRT modality. These measures, aggregated for each information status at the referential and lexical levels of information status, are presented in Table 3.5. and Figure 3.2.

Table 3.5. *Percent positive prominence ratings (%, SD) in the PRT broken down by word information status. Top: lexical information categories; Bottom: referential information categories.*

IS LEX	**l-accessible**	min-max	**l-given**	min-max	**l-new**	min-max
auditory PRT	22.46 (28.4)	0-96.5	15.02 (19.8)	0-89.7	38.3 (29.3)	0-96.6
silent reading PRT	32.27 (25.9)	0-95.7	28.65 (19.73)	0-95.7	53.34 (23.8)	0-95.7

IS REF	**r-bridging**	min-max	**r-given**	min-max	**r-new**	min-max	**r-unused**	min-max
auditory PRT	10.8 (21.8)	0-96.5	17.0 (21.5)	0-89.6	40.9 (28.8)	0-96.6	22.8 (21.3)	0-96.5
silent reading PRT	21.7 (22.9)	0-95.6	30.86 (19.9)	0-95.7	54.66 (23.04)	11.7-95.7	40.3 (19.6)	0-96.5

Figure 3.2. *Percent of words rated as prominent in the PRT broken down by word information status. Left: lexical information categories; Right: referential information categories. Error bars represent standard deviation.*

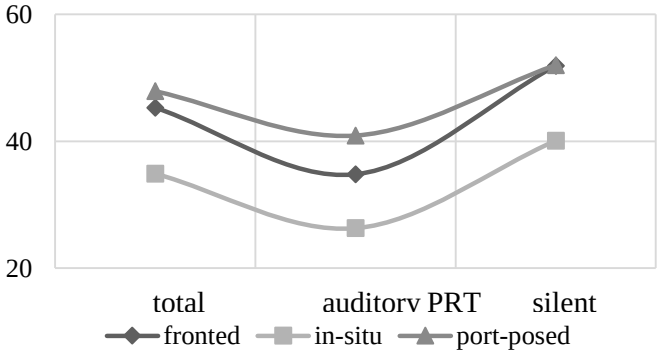


In Table 3.6. and Figure 3.3. shown below, we present percent of respondents who rated in- vs. ex-situ words as prominent, in each PRT modality. Descriptive statistics shown in Tables 3.5. and 3.6. warrant the following observations. First, among the referential and lexical information categories, more raters chose discourse-novel words as prominent, compare to other information categories and regardless of the PRT modality. Second, more raters chose ex-situ words as prominent, compared to in-situ words and regardless of the PRT modality.

Table 3.6. *Rate at which words were identified as prominent (% raters) with reference to their sentence position in the silent reading and auditory PRTs.*

word order	**fronted**	min-max	**in-situ**	min-max	**post-posed**	min-max
auditory PRT	34.8 (23.7)	0-92.8	26.39 (28.4)	0-96.6	40.90 (32.7)	0-93.1
silent reading PRT	51.91 (24.5)	0-95.7	39.37 (25.7)	0-95.6	52.12 (26.7)	0-87.2

Figure 3.3. *Rate at which words were identified as prominent (% raters) with reference to their sentence position in the silent reading and auditory PRTs.*



Individual raters' binary prominence ratings from each PRT version were modeled using two mixed-effects logistic regressions, one per task version. Both models assessed the log likelihood that a word is rated as prominent given its information status, position in the sentence or phrase, and, in the auditory PRT model, its acoustic-prosodic features. Each model featured *p-score* as a binary dependent variable. *Word Order*, *IS_REF*, and *IS_Lex* with levels listed in Table 3.3. were

chosen as predictor variables. The four acoustic-prosodic measures extracted from the model speaker's reading performance entered the auditory PRT model, however, following the examination of the correlation matrix of the fitted coefficients, *max f0* had to be excluded from the auditory PRT regression model due to a strong correlation with *f0 range*. To reflect the possibility that in both PRT modalities raters could identify one or multiple words as prominent per target sefment, as well as account for idiosyncratic trends in the raters' rating styles, the silent reading and the auditory PRT models included random intercepts for *rater* and *word*. The auditory PRT model also included random slopes for *rater* and *mean intensity*, *rater* and *f0 range*, and *rater* and *duration*, to account for individual differences in raters' perception of the model speaker's prosodic expression.

Because of the inherent interrelatedness of the fixed effects in the PRT models, a multi-collinearity check was carried out. A Variance Inflation Factor (VIF) was computed for fitted estimates of the fixed effects. The mean VIF for the silent reading PRT model was computed at 3.37 and ranged between 1.29 and 5.00; the mean VIF for the auditory PRT model was computed at 3.11 and ranged between 1.28 and 6.67. With all obtained VIF measures lower than the threshold level of 10 (Myers 1990), we proceed to reporting the outputs of the p-scores analyses in Tables 3.7. and 3.8.

Results of the Prominence Rating Task

Prosodic effects associated with the intermediate referent categories *r-unused* and *r-bridging* imperfectly pattern with *r-new* information: both these information statuses are associated with greater f0 maxima and mean intensity than *r-given* (*max f0*: t=5.15, p<.001 for *r-accessible* and t=7.25, p<.01 for *r-bridging*; *mean intensity*: t=3.71, p<.001 for *r-accessible* and t=2.65, p<.01 for *r-bridging*) and smaller duration than *r-given* (*r-accessible*: t=3.91, p<.001; *r-bridging*: t=-7.11, p<.001). These results are consistent with Baumann (2006), who reported that intermediate

information categories in German were prosodically accented, however in a more subtle way (L* accent, using ToBi notation), indicative of their semi-active status in discourse.

The silent reading PRT

Consistent with RefLex and relative to the reference category of *l-given* words, *l-new* words were associated with a greater likelihood of being rated as prominent (z=2.14, p<.05). At the referential (*r*-) level of information structure, relative to the category of *r-given* information, the likelihood of being rated as prominent was lower for *r-bridging* words (z=-2.90, p<.005), but greater for *r-novel* words (z=9.31, p<.001), confirming prediction (a). In the silent reading PRT, words positioned ex-situ, specifically, fronted (z=5.98, p<.001) and post-posed (z=3.90, p<.001) relative to the canonical position, had a greater likelihood of being rated as prominent. Raters were equally likely to identify ex-situ post-posed and fronted words as prominent, for a total of 52.12% post-posed words and 51.9% of fronted words, confirming prediction (b).

Table 3.7. *Results of the mixed effects logistic regression (fixed effects only) of the silent reading PRT ratings (Wald Chi2(6)=614.25, p<.001.*

Fixed effects	Levels	Odds ratio	SE	P>| *z* |
Word Order	fronted	.47	.08	<.001
	post-posed	.31	.08	<.001
IS Ref	r-unused	-.21	.9	<.03
	r-bridging	-.24	.08	<.005
	r-new	.87	.09	<.001
IS Lex	l-accessible	.1	.08	n.s
	l-new	.22	.1	.03

The auditory PRT

Consistent with RefLex and relative to the baseline category of *l-given* words, *l-accessible* (z=6.51, p<.001) and *l-new* words (z=5.24, p<.001) were more likely to be rated as prominent. At the

referential (*r*-) level of the information structure, relative to the baseline referent category *r-given*, the likelihood of being rated as prominent was lower for *r-unused* and *r-bridging* referents (z=-2.22, p<.05 and z=-8.68, p<.001, respectively), but greater for *r-novel* referents (z=8.56, p<.001). Results of auditory PRT are in line with results obtained in the silent reading PRT and jointly render prediction (a) confirmed: Regardless of the PRT modality, discourse-new information, at the lexical and referential levels, presents the information category which was more likely to be rated as prominent.

Table 3.8. *Results of the mixed effects logistic regression (fixed effects only) of the auditory PRT ratings (Wald Chi2(6)=1104.65, p<.001.*

Fixed effects	Levels	Odds ratio	SE	P>| *z* |
Word Order	fronted	.35	.07	<.001
	post-posed	.45	.06	<.001
IS Ref	r-unused	-2.27	1.02	.03
	r-bridging	-.81	.09	<.001
	r-new	.75	.09	<.001
IS Lex	l-accessible	.54	.83	<.001
	l-new	.52	.1	<.001
mean intensity		.05	.02	.02
duration		.06	.03	.045
f0 range		.19	.03	<.001

All acoustic-prosodic parameters in the auditory PRT model were positively associated with the likelihood that a word is rated as prominent (f0 range: z=6.48, p<.001; mean intensity: z=2.32, p<.05; duration: z=2.0, p<.05), consistent with prediction (c). Words positioned ex-situ, specifically, fronted (z=5.15, p<.001) and post-posed (z=6.97, p<.001) relative to the canonical position, were also more likely to be rated as prominent, in line with prediction (d). Detailing this finding further, Figure 3.3. demonstrates that raters rated fewer words fronted to the sentence-

initial position, for a total 34.8% fronted words, but more words postposed to the sentence-final position, for a total of 41% of post-posed words, as prominent.

Recall that we documented two types of prosodic variation in the model speaker's read production data, one pertaining to the information structure distinctions in the narratives and the other - to the linearization of sentence constituents. Regarding this second source of acoustic-prosodic variation, we uncovered that in the read production data, ex-situ fronted words underwent consistent prosodic augmentation, while ex-situ post-posed words were partially reduced and had lower pitch maxima and lower mean intensity. To determine if raters equally attended to these distinct sources of prosodic variation, we examined the relationship between the magnitude of acoustic-prosodic parameters estimated in the auditory PRT model and the number of raters who selected a word as prominent, for different information categories, as well as for in- and ex-situ words. Given the lack of consistency in the prosodic expression of the lexical information categories, we focused on the referential information categories best represented in our data, namely, *r-new, r-given,* and *r-bridging*[19]. Results of these analyses are presented in Figures 3.4.-3.6. below.

As Figure 3.4. demonstrates, **mean intensity,** overall, was positively correlated with the number of raters who rated any given word as prominent. However, this relationship did not hold for ex-situ post-posed words, which, as we have previously reported, had overall lower mean intensity in the model speaker's production data (see Table 3.4.). Figure 3.5. demonstrates that **the pitch range** maintained a positive relationship with the number of raters who rated any given word as

[19] Because the PRT results for the *r-unused* information, which is the least represented information type in our data with the most variable p-scores, were overall similar to those obtained for the r-bridging information, to simplify presentation, r-unused was excluded from this analysis.

prominent. This relationship held regardless of the word order, but was less pronounced for *r-bridging* and *r-given* words and more pronounced for *r-new* words.

Figure 3.4. *Prosodic variability in the model speaker's reading performance and % affirmative prominence ratings in the auditory PRT for in- vs. ex-situ words (left panel) and referential information types (right panel). X-axis: Mean intensity (dB); Y-axis: % raters who identified a word as prominent.*

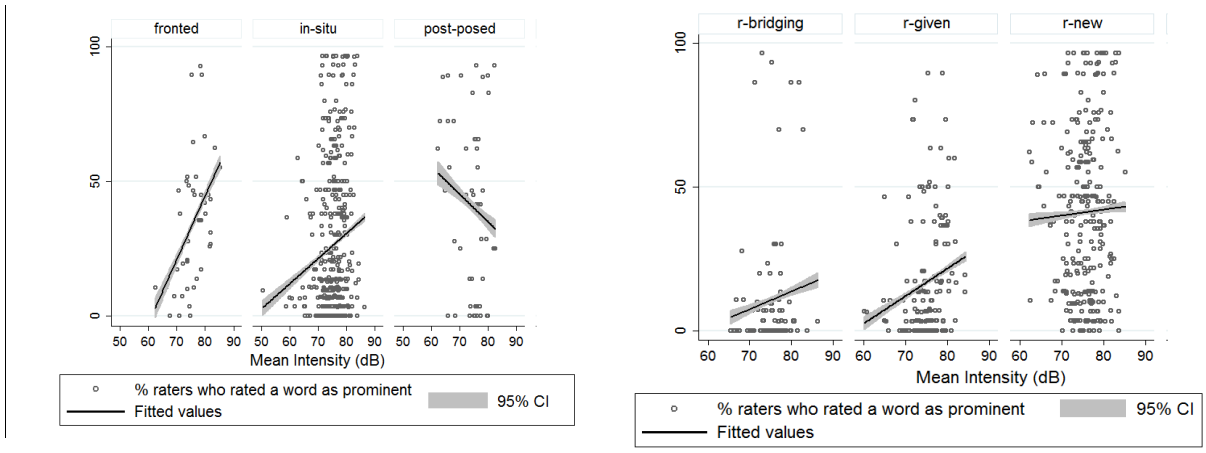


Figure 3.5. *Prosodic variability in the model speaker's reading performance and % affirmative prominence ratings in the auditory PRT for in- vs. ex-situ words (left panel) and referential information types (right panel). X-axis: f0 range (Hz); Y-axis: % raters who identified a word as prominent.*

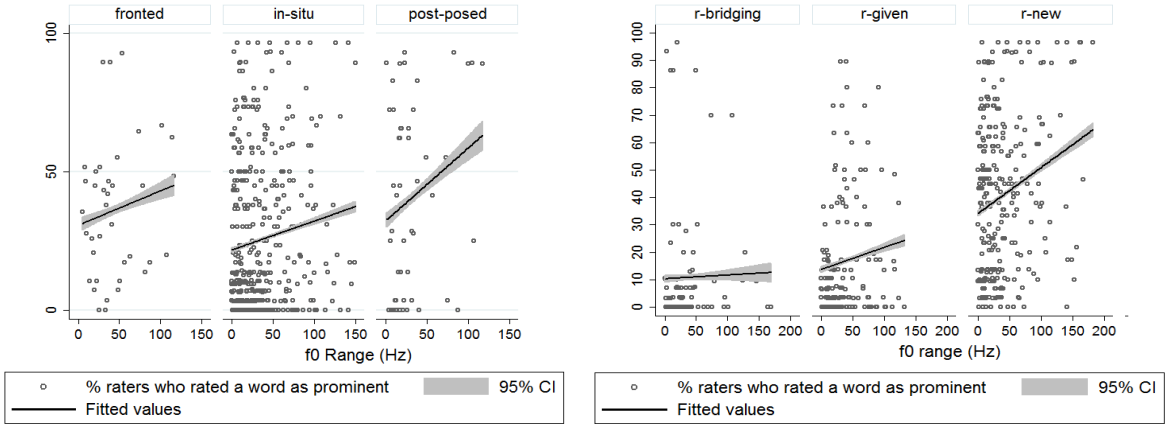
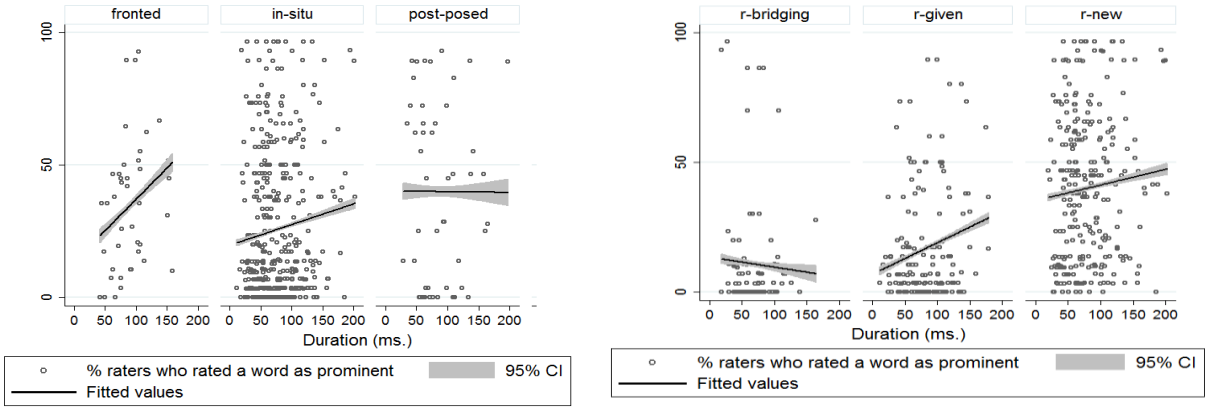


Figure 3.6. demonstrates that **duration**, overall, maintained a positive relationship with the number of raters who rated any given word as prominent. Duration was negatively related to the probability

that an *r-bridging* referent is rated as prominent, in addition to the finding that *r-bridging* information in the model speaker's read production data had reduced duration (see Table 3.4.).

Figure 3.6. *Prosodic variability in the model speaker's reading performance and % affirmative prominence ratings in the auditory PRT for in- vs. ex-situ words (left panel) and referential information types (right panel). X-axis: duration (ms.); Y-axis: % raters who identified a word as prominent.*



In summary, visual examination of Figures 3.4.-3.6. helps us tap into the relation between the acoustic-prosodic variability observed in relation to referential information status and word order. We observed that the information status of a discourse referent not only affects its prosodic characteristics, but also mediates its perceived prominence. This observation is supported by the finding that highly inferable *r-bridging* information was prosodically reduced in the model speaker's reading performance and that percent raters who rated *r-bridging* referents as prominent was proportionate to the acoustic-prosodic profile of the r-bridging information, and is relatively low (mean=10.8% (*SD*=21.8). *R-new* information, on the contrary, was prosodically augmented in the model speaker's reading performance. In the PRT, we observed a positive relationship between the magnitude of the acoustic-prosodic parameters intensity, duration, and pitch, and the number of raters who rated *r-new* referents as prominent (mean=40.9%, *SD*=28.8).

We found that prosodic augmentation specific to ex-situ constituents is in general predictive of how many raters identify them as prominent. This relationship, in particular, holds for acoustic-

prosodic expression of ex-situ fronted words. Notice, however, that for ex-situ post-posed words, the magnitude of acoustic parameters mean intensity and duration is not predictive of the number of respondents who rated these words as prominent. This result points to a possibility that observed acoustic-prosodic changes in the ex-situ post-posed words may be speaker-specific or phonetic in nature. We return to this result in the discussion section.

Relative contribution of word order and prosodic cues to perceived prominence

Auditory PRT data were used to evaluate the relative contribution of acoustic-prosodic and word order cues to perceived prominence in Russian. Our point of departure was the mixed effects logistic model fit to the auditory PRT p-scores (see Table 3.8.). We estimated the goodness of fit for two sets of fixed effects[20] used in this model, as follows. Model fit was gauged using the Bayesian Information Criterion[21] (BIC') for (1) a logistic multivariate model with dependent variable *p-score* and independent variables *IS_REF, IS_LEX,* and *Word Order* and (2) a logistic multivariate model with dependent variable *p-score* and independent variables *IS_REF, IS_LEX, mean intensity, duration, max f0,* and *f0 range*. These models' BIC' values were compared to determine which of the two sets of fixed effects was more likely to have generated the observed data. Goodness of fit test revealed that model (2), in which the p-scores were predicted using the acoustic-prosodic parameters provided a better fit to the data, with the BIC' of -1262.8 (compared to the BIC' of -1231 obtained for the model (1), in which the p-scores were predicted using

[20] Random effects did not enter these analyses as they make it hard to estimate the number of degrees of freedom in the model and obtain legitimate results (Williams 2015).

[21] Unlike pseudo R^2 measures, the information measures for logistic regression models have penalties for including variables that do not significantly improve model fit and can lead to more parsimonious but adequate models when a large data sample is used. When two or more models are compared, the model with the smaller Bayesian Information Criterion coefficient is the one that fits the data best. The difference between the BIC coefficients of 10 or more constitutes 'very strong' evidence for the model with the smaller BIC coefficient (Raftery, 1995).

predictor variable *Word Order* only). Next, we compared the goodness of fit for model (2) and another multivariate logistic model (3) which included *p-score* as dependent variable and *IS_REF, IS_LEX, mean intensity, duration, f0 range,* and *Word Order* as predictor variables. The BIC' of -1308.7 was obtained for this new model proving a substantial improvement over model (2). Comparing the fit of these three models to the auditory PRT data revealed that acoustic-prosodic cues alone explain more variability in the auditory PRT p-scores than word order alone. However, it is the combination of word order and acoustic-prosodic cues which yielded the best fitting model for the auditory PRT data.

Summary of the PRT results

Per default prediction (a), we anticipated the information status of a word to scale its perceived prominence. Results of the regression models fit to the silent reading and auditory PRT p-scores corroborate this prediction. In both PRT modalities, discourse-novel information, at referential and lexical levels, was more likely to be rated as prominent. At the referential level, highly inferable *r-bridging* information had the least likelihood of being rated as prominent. Less straightforward results were obtained for the intermediate category of lexically-accessible information. In the auditory PRT, *l-accessible* information patterned with *l-new* information, and was more likely to be rated as prominent than *l-given* information. We attribute this result to the fact that words coded as *l-accessible* had no prior mentions in the narratives and could be judged as discourse-new by the raters who completed the auditory PRT. Recall however, that such words were not prosodically augmented in the model speaker's reading performance and were not associated with a greater likelihood of being identified as prominent in the silent reading PRT.

The result that discourse-novel information in both PRTs was rated as more prominent than discourse-given information is highly consistent with what has been previously reported for the information status distinctions across languages and validates the PRT experimental methodology. Next, per predictions (b) and (d), we anticipated ex-situ words in the narratives to have greater perceived prominence in the silent and auditory PRTs, respectively. These predictions were fully borne out in both PRT versions. Per prediction (c), we anticipated raters to identify words which were prosodically augmented in the model speaker's reading performance as prominent and expected to observe a positive relationship between the magnitude of an acoustic-prosodic parameter and the likelihood that any given word is rated as prominent. This prediction was borne out, as well: in the auditory PRT, prosodic augmentation in all analyzed parameters was indeed positively associated with the likelihood that a word is identified as prominent.

In summary, our experimental results provide evidence that in Russian, prosodic properties of a word and its position in a phrase relative to the canonical constituent order, via an association with its information status, contribute to its perceived prominence. Results of the silent reading PRT are consistent with Hypothesis 3 that an ex-situ position serves as an independent cue to perceived prominence, even when no speaker-generated prosody is supplied. Results of the auditory PRT further suggest that in the spoken modality, listeners are aware of the linearization irregularities in discourse and treat ex-situ position as a cue to prominence. As a result, structural prominence factors into perceived prominence, on a par with acoustic-prosodic augmentation.

We found that acoustic parameters mean intensity, duration, and pitch range exhibited predictable variation whenever a word appeared ex-situ. For example, fronting a word resulted in augmenting its acoustic-prosodic characteristics, regardless of its information status. Post-posing a word, on the contrary, triggered reduction in the prosodic parameters mean intensity and pitch range in

combination with lengthening. In our earlier work, we documented similar patterns of acoustic-prosodic dynamics in read production data from fifteen Russian speakers (Luchkina and Cole 2016). We proposed that these patterns may be, at least in part, explained by the proximity of the fronted and post-posed ex-situ positions examined to the clausal boundaries, which in Russian coincide with the intonation phrase boundaries. A boundary-adjacent position may contribute to strengthened acoustic-prosodic expression of the clause-initial word and reduced acoustic-prosodic expression of the clause-final word. We will discuss word order-specific patterns of acoustic-prosodic variation in more detail in the following section.

Estimations of goodness of fit for fixed effects used in statistical modeling of the auditory p-scores lead us to conclude that acoustic-prosodic cues examined in this study account for more variation in the prominence ratings than the linearization cue. This result is consistent with the observation that non-canonical word orders, while not infrequent in Russian, present a relatively restricted, and therefore less common prominence marking cue. Prosodic expression, on the other hand, is available to the speaker regardless of the word order and is contributed to by multiple acoustic parameters which are jointly used to deliver a perceptibly more enhanced rendition of the prominent word.

Discussion

The goal of this work is to parameterize perceived prominence in a free word order language, such as Russian, and understand which factors guide naïve readers' or listeners' perception of a word as prominent in discourse or narrative. To answer these questions, we tested whether variation in word order and acoustic-prosodic expression can serve as means of expressing the information status of a word and, by doing so, mediate its perceived prominence. We then gauged the relative contribution of word order and acoustic-prosodic cues to perceived prominence in Russian. We

worked with discourse samples which come from two published narratives. These narratives successfully capture the interrelatedness of word order and prosodic accenting as means of expressing relative information accessibility and perceived prominence.

Information status distinctions predict perceived prominence. Following Bornkessel and Schlesewsky (2009), Baumann and Riester (2013), and Arnold et al. (2000), we deployed gradient differences in information givenness and accessibility as a prominence scale. We used the information structure annotation scheme proposed by Baumann and Riester (2012, 2013) as the basis for discriminating between qualitatively different categories of information at two distinct levels, referential, pertaining to a word referent, and lexical, pertaining to the lexical item expressing that referent. Our default prediction was that under a successful implementation of the unguided prominence rating task (PRT), we would observe that the likelihood that a word is rated as prominent is commensurate with its information status in discourse. Consistent with the earlier work on information structure and its relation to prominence (Breen et al. 2010, Watson 2010), we anticipated discourse-given information, at the lexical and referential levels, to be the least prominent information category. We anticipated r-novel and l-novel information categories, least accessibly to the reader or the listener, to be more prominent compared to discourse-given information. Despite the fact that no definition of perceived prominence was offered to linguistically naïve raters, and regardless of the PRT modality, lexical words and referents new to discourse were more likely to be rated as prominent. In contrast, an intermediate referential information category ‘r-bridging’ was associated with a lower likelihood of being rated as prominent, consistent with our prediction (see Tables 3.7. and 3.8.). Prominence ratings obtained for referential information categories confirmed that distinctions in information accessibility and givenness at the referential level present a legitimate prominence scale.

RefLex differs from other information structure annotation schemes in that it discriminates between a number of lexical information categories, focusing on the newness of a lexical word and blind to the information status of the word referent. In this study, lexical information categories yielded less clear results with respect to perceived prominence: Raters did not discriminate between words which were lexically new (l-new) and words which were lexically-new but had a high cloze probability (l-accessible). Furthermore, both these lexical information categories were more likely to be rated as *less* prominent than l-given words, regardless of their prosodic characteristics. Riester and Piontek (2015) argue that given words may be (pitch)-accented, especially when focused or under the influence of extra-linguistic factors, such as individual speaker decisions. However, they do not specify if the kind of givenness they discuss pertains to the word referent or the lexical word. We conclude that lexical information status failed to yield a coherent prominence scale in this study.

Next, we focused on the relative contribution to perceived prominence of two distinct mechanisms which present known cues of information status in Russian, word order and prosody.

Prosodic correlates of perceived prominence. Acoustic-prosodic variability presents one major way of communicating pragmatic meaning and prominence in discourse, cross-linguistically. In languages with limited word order freedom, like English, via the mechanism of prosodic prominence displacement, the nuclear pitch accent aligns with the most prominent word in a phrase or clause (Calhoun 2010) As a result, the prominent word undergoes prosodic augmentation, irrespective of the linear order of the sentence constituents. Under free word order, prominence displacement may be obviated by re-ordering of the sentence constituents; however, this alternative prominence marking strategy may be considered secondary to the use of acoustic-prosodic cues due to higher cognitive costs associated with constituent reordering (Skopeteas and

Fanselow 2012).

In this study, we mostly focused on three distinct acoustic parameters, including the range of the fundamental frequency or pitch (f0 range), mean intensity and duration. These acoustic parameters were chosen because they represent important psycho-acoustic correlates of prosodic prominence, pitch, segment loudness and segment length. Our findings reveal that all of the acoustic parameters selected for this study serve as cues to the information status of a word, its position in a sentence or phrase relative to the canonical SVO order and the likelihood that the listener perceives that word as prominent. In this respect, Russian is not different from other languages which use prosodic distinctions in spoken discourse.

Our findings confirm that acoustic-prosodic variability in the Russian language is used in ways which are similar to and different from those documented for a fixed word order intonational language like English. On the one hand, acoustic-prosodic cues in Russian reliably mark information status of a word, and the magnitude of acoustic measures is positively correlated with the likelihood that a word is perceived as prominent. On the other hand, unlike in fixed word order languages, selective augmentation of acoustic-prosodic parameters also marks words which appear in non-canonical positions in a sentence or phrase.

In order to obtain a more detailed picture of acoustic-prosodic variation in relation to the information status of a lexical word and its referent, consistent with RefLex information structure annotation scheme (Baumann & Riester 2012, 2013), we examined prosodic profiles of words of different information status in two stylistically varied discourse samples. Lexical information categories, in general, generated less clear results with respect to their acoustic-prosodic profiles. Compared to the baseline category of l-given words, l-new words in the model speaker's production were *not* augmented, but instead had reduced duration and pitch range. Unlike lexical

information status, referential information status scaled acoustic-prosodic parameters intensity, duration, and pitch range in the direction inverse to referent accessibility, leading discourse-novel referents in the corpus to have augmented prosodic expression. Relatively more accessible discourse-inferable information appeared de-accented at referential and lexical levels of annotation, as evident from significantly reduced intensity, pitch range, and maximum pitch for l-accessible information, and reduced duration for r-bridging information.

Analyses of perceived prominence ratings revealed that acoustic-prosodic variability specific to the word information status, is not limited to production, but is also actively deployed by the listener during discourse comprehension. To illustrate, in the auditory PRT, raters treated the acoustic-prosodic realization of a word as a cue to its discourse status and relative prominence. The likelihood that a word was perceived as prominent was positively correlated with augmentation of acoustic-prosodic parameters pitch range, duration and vowel intensity. In the auditory PRT, such likelihood was lower for prosodically reduced information categories, including highly inferable r-bridging information and fully accessible r-given information. The likelihood of being rated as prominent was greatest for the least accessible r-new information, which featured augmented duration, pitch and intensity in the model speaker's reading performance. These results confirm that prosodic augmentation is inversely related to referent accessibility in spoken discourse but maintains a positive relationship with the likelihood that a word is perceived as prominent.

Prosodic variation in the model speaker's read production reflected not only of information accessibility but also argument linearization in a sentence or phrase. Acoustic-prosodic parameters duration, mean intensity, and pitch range varied systematically when a word occurred ex-situ in a

sentence or phrase, relative to the canonical SVO order. We discuss word order-related prosodic variability and its contribution to perceived prominence in more detail in the subsequent sections.

Structural correlates of perceived prominence. Approximately 20% of the utterances in discourse samples which we selected for the PRT experiment deviate from the canonical SVO word order. This allowed us to test whether surface constituent order affects perceived prominence in Russian. We focused on two distinct constituent linearization patterns, whereby a word appears ex-situ and is fronted or post-posed to a sentence-peripheral position.

Per our research hypothesis, under free word order, positioning a word ex-situ acts as an independent cue to prominence, which may be further reinforced with acoustic-prosodic features exclusive to such a position. This hypothesis was supported with two different types of evidence obtained from the PRT materials. The first of them is comprised of distributional properties of ex-situ words in the PRT discourse samples. As Figure 3.1. demonstrates, given information in the PRT materials has a tendency to occur in-situ or appear fronted, in other words, occur relatively early in a sentence or phrase. Novel information, on the contrary, has the most ex-situ occurrences and favors the pitch-accented sentence-final position, thereby conforming with the 'given before new' information template preferred in the Russian language (Neeleman and Titov 2009, Slioussar 2011b, Ionin and Luchkina, under review). Such information structure template has been characterized as facilitative for both the speaker and the listener/reader, cross-linguistically, (Arnold et al. 2000, Branigan 2007, Clifton and Frazier 2004), as it allows the speaker to retrieve highly accessible, already active words and referents before the novel, less accessible ones optimizing production and comprehension processes. By easily monitoring the information structure in discourse, the listener/reader may better gauge the relative contribution or prominence of a word to discourse meaning.

The second type of evidence supporting our hypothesis concerns a special prominence status of the ex-situ words in the PRT materials, and is in line with the prediction that an ex-situ position acts as an independent cue to perceived prominence. Results of the silent reading and auditory PRTs confirmed that regardless of the direction of surface constituent movement or the prosodic characteristics of the ex-situ constituents, they were more likely to be perceived as prominent.

Cross-application of structural and acoustic-prosodic variability in Russian. The cross-application of prosodic and word order variability in Russian, which makes spoken discourse a multi-cue environment, is of special interest for this study.

In our earlier work, we used read production data from fifteen Russian speakers to compare the acoustic-prosodic expression of in- and ex-situ words located at phrasal boundaries (Luchkina and Cole 2016). We found that regardless of whether the ex-situ word was located phrase-initially or phrase-finally, one or more of its acoustic parameters were augmented significantly more than in the control sample of the in-situ words located at phrasal boundaries. For example, the sentence-final lengthening of the ex-situ post-posed words was significantly greater than that of their in-situ counterparts. Consistent with this result, in the present study, prosodic properties of the ex-situ words in the model speaker's read production overall reflected whether the word occurred phrase-initially or phrase-finally.

A closer look at the acoustic-prosodic make up of the ex-situ words in PRT materials revealed that regardless of the surface movement type, they featured augmented duration; other acoustic-prosodic changes included reduction in mean intensity and pitch range for post-posed, phrase-final ex-situ words, but systematic increase in these parameters for fronted, phrase-initial words. These results, in part, are in line with a series of earlier studies on special acoustic-prosodic status of ex-situ constituents in Hindi (Patil et al. 2008, Luchkina, Puri, Jyothi, Cole 2015), Romani (Armaniti

and Adamou 2011), Greek (Baltazani 2003), and Finnish (Vainio and Järvikivi 2006). Systematic prosodic augmentation of ex-situ constituents has been reported for all these languages. In this study, we have expanded this finding further, by showing that in a free word order language like Russian, surface movement to the left or right phrasal boundary triggers qualitatively different changes in the acoustic-prosodic expression of ex-situ words, including a comprehensive prosodic augmentation at the left phrasal boundary and a partial prosodic reduction at the right boundary.

An observation that phonetic realization of individual segments is affected by their position in the prosodic structure, and, specifically, by the proximity to a prosodic domain boundary, was previously offered in work by Keating, Cho, Fougeron, Hsu (1993) and Pierrehumbert and Talkin (1992). These authors experimentally showed that domain-initial segments have greater duration and 'greater articulatory magnitude' leading to prosodic strengthening at the phone, syllable, or word level (Keating et al. 1993, Cho 2011). Results reported by Keating et al. prompt that reduction in acoustic parameters intensity and pitch range that we documented for phrase-final ex-situ words may present a case of articulatory weakening occurring at the right boundary of an intonation phrase.

Proximity of an ex-situ constituent to a phrasal boundary provides one possible explanation for the distinctive acoustic-prosodic realization of ex-situ words in the read production data presented in this study and accounts for articulatory strengthening of the fronted words, as well as a partial acoustic-prosodic reduction of the post-posed words. The auditory PRT results revealed that prosodic augmentation of ex-situ fronted words was interpreted by the listeners as a prominence lending cue, and factored into the prominence scores associated with these words (see Figures 3.4.-3.6.). At the same time, during the auditory PRT, more phrase-final (as opposed to phrase-initial) ex-situ words were rated as prominent, despite their partial prosodic reduction (see Figure 3.3. and

Table 3.5.). We propose that reduction in mean intensity and pitch range observed at the right phrasal boundary could be offset by important phonological status associated with this position. As in other right branching languages, in Russian, the rightmost constituent of an intonation phrase has to be its head. This means that by post-posing a word to the rightmost intonation phrase boundary, the speaker makes it the nuclear pitch-accented head of the intonation phrase and increases the likelihood that it is perceived as prominent. We reason that during auditory comprehension, the nuclear pitch accent could serve as a compensatory perceptual trade-off for the reduction in intensity and pitch range in the post-posed words. Ample evidence showing that listeners in intonation languages are biased to hear nuclear pitch accented words as prominent, even when the acoustic-prosodic cues supporting this perceptual experience may not be (fully) available, supports this view (Hermes and Rump 1994, Terken and Hermes 2000, Calhoun 2010).

Reflecting on the production aspect of non-canonical orders, we add that acoustic-prosodic strengthening characteristic of ex-situ words may be due to a more effortful production associated with the non-canonical orders in some free word order languages, including Russian. Following Slioussar (2011b), Sekerina (2003), Dragoy and Baastiance (2009), non-canonically ordered utterances in Russian are computationally more complex, to the extent that they may be unavailable to vulnerable speaker populations, including young children and aphasiac patients. Consistent with this account, acoustic augmentation of the ex-situ words in our study may, in part, relate to a strengthened articulatory effort serving as a reflection of the greater cognitive effort that production of the non-canonical orders involves.

Both boundary-adjacent strengthening (weakening) and scrambling complexity accounts predict that the acoustic parameters of ex-situ words may vary orthogonally to their prominence status. Our findings are consistent with this prediction. We have demonstrated that the acoustic-prosodic

characteristics of the ex-situ words in our data are imperfectly correlated with the likelihood that they are perceived as prominent. Specifically, a partial prosodic reduction of the post-posed words did not affect their perceived prominence ratings. Recall that in both PRT versions, ex-situ fronted and ex-situ post-posed words had a greater likelihood of being perceived as prominent than their in-situ counterparts. Furthermore, fronted and post-posed words were rated as prominent at a comparable rate in the silent reading PRT. In the auditory PRT, contrary to the decline in intensity and pitch range observed in ex-situ post-posed words, they were rated as prominent *more often* than the fronted words. These results confirm that acoustic characteristics of the ex-situ post-posed words in the model speaker's read production varied irrespective of their information status and perceived prominence. We conclude that listeners are aware of the nature of acoustic-prosodic variability coincidental with a change in word order in Russian and probabilistically use acoustic-prosodic expression of the ex-situ words as a cue to their discourse status and relative prominence. When the acoustic-prosodic cues are predictive of the perceived prominence of a word, they add up to the likelihood that that word is perceived as prominent. However, when the acoustic-prosodic parameters vary orthogonally to the word's perceived prominence, they lose their perceptual significance and get diminished to redundant phonetic dynamics in speech, possibly offset by the inherent perceptual salience of the nuclear pitch accented phrasal position. This conclusion lends support to Féry (2013)'s proposal that augmented acoustic-prosodic expression is not a prerequisite for prominence. Rather, prosodic strengthening of prominent information corroborates a prominent reading of a word when available and may further reinforce a structural prominence cue, such as proximity to a phrasal boundary. Further support for this view comes from a study by Vainio and Järvikivi (2006), who presented compelling experimental evidence pointing to a special perceptual salience of ex-situ words in Finnish, regardless of their prosodic expression. Vainio and

Järvikivi conducted a series of auditory perception experiments in which magnitude of pitch peaks (height and slope) and intensity varied in stimuli with neutral (canonical) and altered (non-canonical) word order. Finnish listeners were more prone to perceive ex-situ words as prominent regardless of the controlled prosodic manipulation. Vainio and Järvikivi proposed that an ex-situ position has a robust top-down influence on the perception of the main sentence stress and drives the interpretation of the ex-situ word as prominent, even when the prosodic cues associated with the ex-situ position are made unavailable to the listener.

This study reveals that the cue validity of prosodic parameters may be lower for ex-situ words, supplanted with their high positional or structural saliency. Despite this finding, the important contribution of the acoustic-prosodic cues to perceived prominence in Russian should not be underestimated. Results of the auditory PRT, during which acoustic-prosodic information and constituent linearization were simultaneously available to the listeners, allowed us to gauge the relative contribution of each cue type to perceived prominence in Russian. Comparing model fit for three different combinations of p-score predictors, we determined that acoustic-prosodic parameters explain more variance in the dependent variable (auditory) p-score than word order. We attribute this result to the fact that in the PRT materials, word order variability involved about 20% of the sentences, whereas prosodic variability was available for every word and presented a more readily available channel of information about the word's discourse status and relative prominence. We further determined that the optimal modeling of perceived prominence is achieved when acoustic-prosodic parameters *and* word order were jointly used as predictors of the likelihood that a word is perceived as prominent. This result supports the conclusion that prominence in Russian is a product of information linearization preferences and acoustic-prosodic

augmentation which apply separately or in combination to cue the relative prominence of a word in discourse.

Conclusion

This study contributes to the understanding of perceived prominence in free word order language and the linguistic mechanisms which render a word perceptibly prominent in read discourse. We report that in Russian, a free word order language, perceived prominence reflects the information status of a word and may be expressed be means of prosody or through a combination of acoustic-prosodic and constituent linearization processes.

Analysis of word-level perceived prominence ratings provided by linguistically naïve native speakers of Russian revealed that non-canonical constituent linearization serves as a source of structural prominence in Russian. By aligning a word with an inherently prominent position in a sentence or phrase, such as left or right (intonation) phrase boundary, the speaker renders that word prominent. In the more numerous cases, in which information status and perceived prominence are expressed in-situ, Russian is highly similar to languages with rigid word order like English, in that it relies on the use of acoustic-prosodic cues to prominence. In such cases, prosodic augmentation serves as a key prominence lending cue. As canonically ordered clauses account for 80% of the materials used in this study, prosodic cues constitute the main channel for expressing perceived prominence and motivate a strong positive relationship between the magnitude of acoustic-prosodic parameters intensity, duration, and pitch, and the likelihood that a word is perceived as prominent. We have shown that the magnitude of these acoustic-prosodic parameters depends on the information status of a word and that word's phrasal position. These factors determine, probabilistically, whether the listener relies on the word's prosodic expression as a cue to its relative prominence or prioritizes positional characteristics of the ex-situ word. Hence, while

prosodic cues may be more readily available during auditory discourse comprehension, it is a combination of acoustic-prosodic and structural variability that explains why some words are perceived as more prominent than others best.

Questions for future research

In this study, we empirically show that in Russian, as in languages with fixed word order, perceived prominence is closely related to the information status of a word. Unlike in fixed word order languages, Russian speakers use prosodic cues and word order to mediate the visibility of a word in discourse and signal its relative prominence. In the silent reading prominence rating task, raters reliably discriminated between words of different information status and paid attention to the linear order of sentence constituents as a reflection of their information status and relative prominence. In the auditory modality PRT, perceived prominence ratings continued to robustly reflect information status distinctions expressed through word order and prosodic variability in discourse. These findings point to an inherent, modality-independent ability of the reader and the listener to closely monitor information structure in discourse and present new challenges to understanding perceived prominence in a multi-cue environment such as a free word order language discourse.

One of these challenges lies in understanding the motivation behind prosodic augmentation of ex-situ words which we documented for Russian and which had been previously documented for a number of free word order languages. As discussed in the preceding sections, acoustic-prosodic make up of ex-situ words may present a phonetic epiphenomenon of articulatory strengthening/weakening at clause-peripheral positions. Alternatively, it can point to a greater cognitive effort that production of non-canonical orders involves, regardless of the underlying properties of the ex-situ constituents or the extra-linguistic meaning conveyed by non-canonically ordered utterances (Bard and Aylett 1999, Bard, Andreson, Sotillo, Aylett, Doherty-Sneddon,

Newlands 2000, Pluymaekers, Ernestus, Baayen, 2005). We recommend that future studies conduct controlled experimental investigations to evaluate these proposals.

In this study, we focused on the use of acoustic-prosodic variability and word order as correlates of perceived prominence in Russian. Our findings lend partial support to the hypothesis that acoustic-prosodic cues and the linear order of sentence constituents reinforce each other in signaling the relative perceived prominence of a word. We have observed that prosodic augmentation systematically marks ex-situ words in the sentence-initial position, however the augmentation of sentence-final words is partial, and engages segment duration only. Despite the selective nature of acoustic-prosodic changes, fronted and post-posed ex-situ words in our materials had a greater likelihood of being perceived as prominent. The observed selective application of acoustic-prosodic cues in fronted and post-posed words motivates the question about the inherent perceptual salience of a non-canonical position in Russian and other free word order languages. This question has been previously addressed by Vainio and Järvikivi (2006) who found that in Finnish, perception of ex-situ words is top-down and expectation-driven rather than based on speaker-specific prosodic choices. Results of the present study further motivate testing the inherent perceptual salience of ex-situ words in typologically different constituent dislocating languages, such as Russian, Hungarian, and Hindi.

CHAPTER 4

PROCESSING OF NON-CANONICAL ORDERS IN READ DISCOURSE IN RUSSIAN (Paper 3)

Abstract

Augmented prosodic expression of a sentence constituent and a change in word order may co-occur in spoken and read discourse in a number of free word order languages, including Russian. A number of studies (e.g., Fraundorf, Watson, and Benjamin 2010) demonstrated that prosodic augmentation, via pitch accenting, leads to a better and more accurate recall of information relayed by the accented word. At the same time, non-canonically ordered sentences have been characterized as computationally more resource-intensive, associated with longer reading times, and repetition disfluencies (e.g., Kaiser and Trueswell 2004, Dragoy and Baastiance 2010).

This study examines whether prosodic augmentation of ex-situ words in Russian, a free word order language, is conducive to auditory sentence processing and comprehension and leads to faster recognition of information relayed by the ex-situ word. Results of a lexical probe recognition task completed by linguistically-naïve native speakers of Russian are reported. Consistent with earlier work on Russian (Sekerina 1999, Slioussar 2011b), this study finds that native Russian speakers take significantly longer to identify a lexical probe originally presented in a non-canonical OVS order. The cross-application of non-canonical constituent order and prosodic augmentation of the ex-situ word facilitates subsequent recognition of a lexical probe but only when the ex-situ word is aligned with in a natural prominence landing site, such as the nuclear pitch accented sentence-final position in Russian.

Introduction

Word order variability is used in morphologically rich languages to express information status of discourse entities and relay their relative prominence in discourse (Calhoun 2010). There has been a consensus in psycholinguistic literature (see Sekerina 1999 for a review) that processing and production of non-canonical orders is context-restricted and resource-intensive, to the extent that constituent reordering precludes other ways of expressing pragmatic focus and information prominence, such as prosody. Indeed, in some free word order languages, making an ex-situ constituent prosodically prominent is not felicitous (e.g., Italian, see Swerts et al. 2002). However, in a number of languages, including Georgian, Greek, and Finnish, prominent words feature augmented prosody independently of their position in a sentence or clause (Skopeteas and Fanselow 2010, Baltazani 2010, Vainio and Järvikivi 2006).

In the absence of prior experimental work, it is unclear whether positioning a prominent word ex-situ and augmenting its prosodic expression results in a yet higher processing load or, on the contrary, facilitates sentence processing. In this study, this issue is explored for Russian, a highly free word order language. This study asks, first, how a change in word order affects the acoustic-prosodic expression of an ex-situ word in Russian. Following earlier work on Russian carried out by Botins et al. 2005 and Luchkina and Cole (forthcoming), the focus is on the parameters f0, intensity, and duration, which contribute to the acoustic-prosodic expression of an ex-situ word. Next, this study asks whether augmented prosodic expression of ex-situ sentence constituents is deployed by listeners during auditory comprehension and if so, whether it has an effect on subsequent recognition of the ex-situ word.

One pertinent observation regarding word order variability in Russian concerns the high computational load associated with production and perception of non-canonical orders. While

discourse-conditioned reordering of sentence constituents is often regarded as *surface* constituent movement and does not involve the level of deep (syntactic) structure, a number of psycholinguistic studies have focused on the processing aspects of non-canonical orders in various free word order languages (Featherston 1999, Clahsen and Fetherston 1999, Nakayama 1995, Kaiser and Trueswell 2004). Among these studies are Sekerina (1999) and Slioussar (2011a) who investigated processing of non-canonical orders in Russian. Both studies reported an increase in silent reading times for non-SVO sentences presented out of context, which they attributed to a more complex syntactic derivation associated with non-canonical orders in Russian. Further, Dragoy and Baastiance (2009) presented experimental evidence that elicited production of non-canonically ordered utterances by Russian-speaking aphasia patients, likewise, is effortful and error-prone.

To summarize, in Russian, as in other free word order languages, word order variability has been tied to information structure and pragmatic focus and may present an effective tool of cueing relative information prominence. Two interesting observations related to word order variability in Russian have been offered in the prior research: non-canonically positioned words have been reported to feature augmented acoustic-prosodic expression, while non-canonically ordered sentences have been linked to added computational costs during production and comprehension.

The listener's prospective

To date, it remains largely unclear if listeners deploy word order, prosody or both as cues to pragmatic focus and relative information prominence. It is possible that the augmented prosodic expression of ex-situ words emanates from (greater) articulatory effort due to a non-canonical linearization of sentence constituents. In situations in which the ex-situ constituent appears sentence-initially or sentence-finally, its acoustic-prosodic expression may further be conditioned

by the immediate proximity of an ex-situ word to a prosodic boundary (Cutler, Dahan, Van Donselaar 1997). If greater computational complexity and/or adjacency to a prosodic boundary have an effect on the acoustic-prosodic expression of ex-situ words, the latter should be orthogonal to the propositional content or discourse status of such words and should be disregarded by the listener. Instead, the linear constituent order should be used as a cue for discourse status and relative information prominence.

On a different account, acoustic-prosodic expression and constituent linearization may be viewed as integral components of a unitary prominence encoding mechanism in a free word order language like Russian, whereby word order flexibility may be viewed as a vehicle used to align the focused word with a designated sentence position where its discourse status is manifest. Indirect empirical support for this account comes from work on Finnish, a highly free word order language, in which a change in word order alters the intensity profile and the tonal shape of a non-canonically ordered utterance (Vainio and Järvikivi 2006, 2007). In a controlled perception experiment, Vainio and Järvikivi (2007) presented Finnish listeners with sentences in which intensity and f0 measures for the ex-situ words were manipulated such that they were within the range associated with their in-situ counterparts. Despite this experimental manipulation, Finnish listeners reported hearing ex-situ words as prosodically prominent. This finding lead Vainio and Järvikivi to conclude that perception of ex-situ words in Finnish is expectation-based and top-down, and informed by the underlying characteristics of the designated sentence positions to which they move.

The present study

This study explores how prosody and word order function independently and in combination during auditory comprehension of discourse in Russian. The first goal of the experimental tasks reported below is to obtain confirmatory evidence that in Russian, a change in word order has an

effect on the acoustic-prosodic expression of ex-situ words. The second goal is to establish if the cross-application of prosodic augmentation and a change in word is functionally meaningful, and as such facilitates subsequent recognition of the ex-situ word.

For the purposes of the present investigation, word order and prosodic properties of 16 canonical SVO and 16 non-canonical OVS sentences in Russian were manipulated. The focus of the read production data analyses reported in the following sections is the acoustic-prosodic characteristics of ex-situ words in the experimental materials, including segment duration, intensity, and tonal characteristics. In the Probe Recognition task reported in Section 2.2., the relative contribution of acoustic-prosodic augmentation coincidental with a change in word order in Russian is gauged to the subsequent recognition of lexical probes matching the ex-situ word.

Production task

Materials and method

Sixteen canonical SVO sentences in Russian each featuring an animate subject and an inanimate object were re-written as object-initial and subject-final OVS sentences. Most (>90%) subject and object nouns in these sentences were unambiguously marked for morphological case, Nominative vs. Accusative, which additionally disambiguated the syntactic roles of the noun phrases in the OVS sentences. For the sentences with ambiguous case marking (the object noun has a null Accusative ending, which makes it morphologically similar to the subject noun), extra care was taken to disambiguate the grammatical functions of these noun phrases by choosing semantically transparent agent-patient pairs. In each resulting SVO-OVS sentence pair, a target noun was identified which was the subject noun or the object noun of the sentence. The two levels of the factor *Word Order* (target noun in- or ex-situ) were then crossed with two levels of the factor *Prosodic Prominence* (target noun prosodically neutral or prosodically prominent).

Two recordings of each experimental sentence were produced by a native Russian speaker, female, age 24. Recording (a) had neutral intonation; in recording (b), the target noun was prosodically prominent, as shown in the spectrograms displayed in (4.1). To minimize the differences between the recorded stimuli, recording (a) was used as the basis for the splice (c). In splice (c), the target noun in recording (a) was replaced with its prosodically prominent counterpart from recording (b), as shown in (4.2) below. Recordings were made in a sound-proof booth using a Marantz PDM 750 solid state recorder and a head-mounted microphone.

(4.1.) Sample stimulus:

Devochki	slushayut	muzyku
girls-NOM	are listening	music-ACC

Girls are listening to music.

Figure 4.1. *Spectrogram of the stimulus in (4.1), canonical order and neutral prosody*

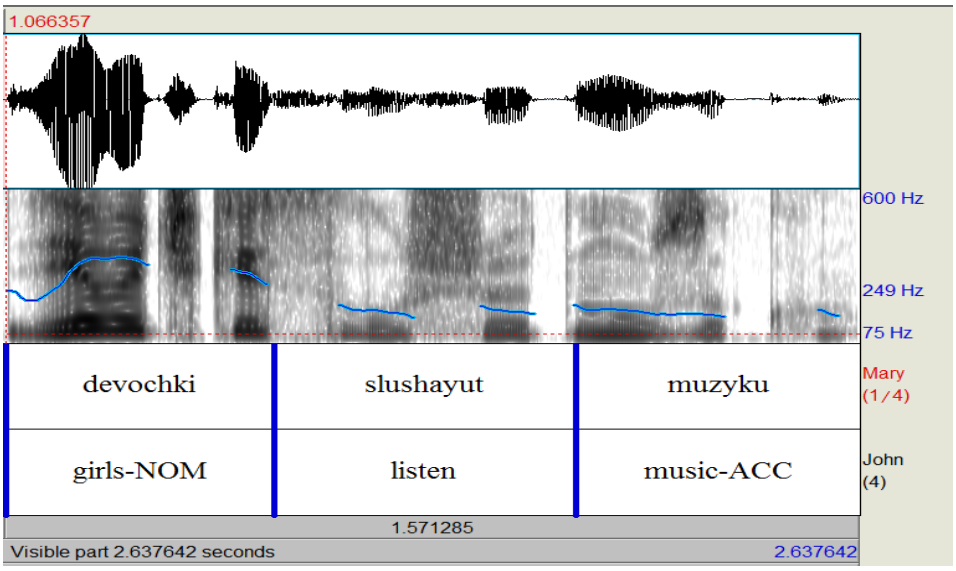


Figure 4.2. *Spectrogram of the stimulus in (4.1), canonical order, sentence-final object NP is prominent*

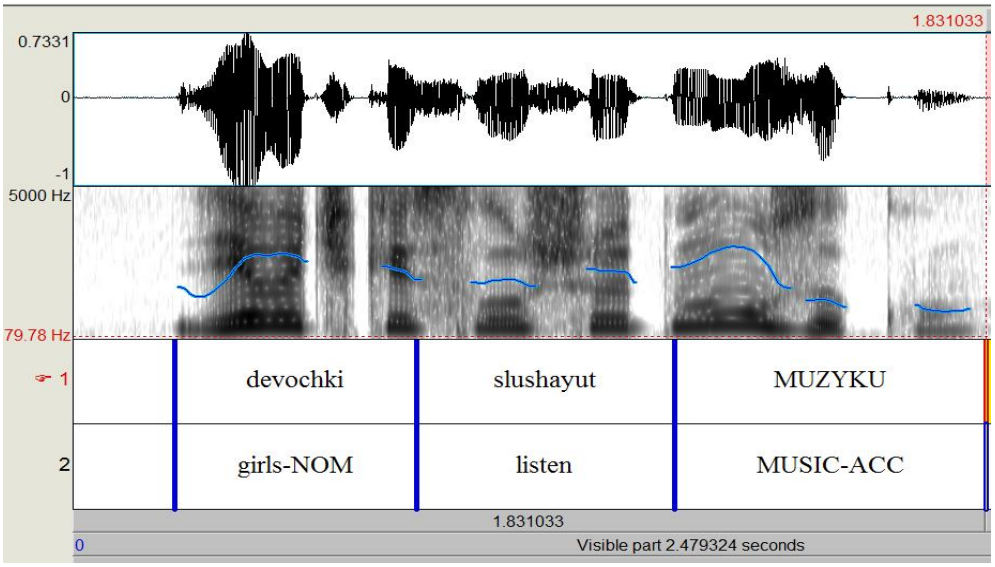


Figure 4.3. *Spectrogram of the stimulus in (4.1), non-canonical order, sentence-initial object NP is prominent*

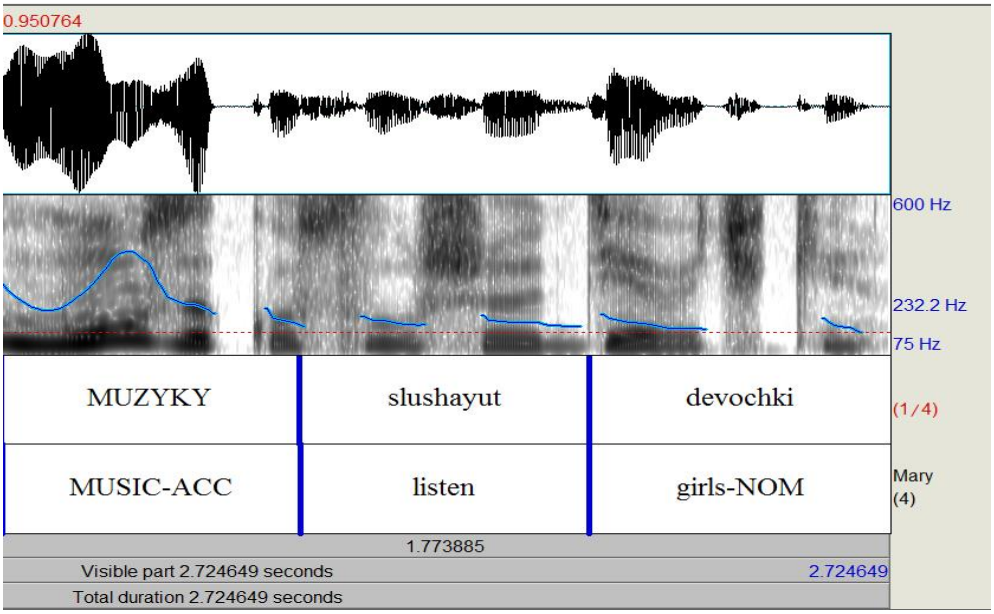
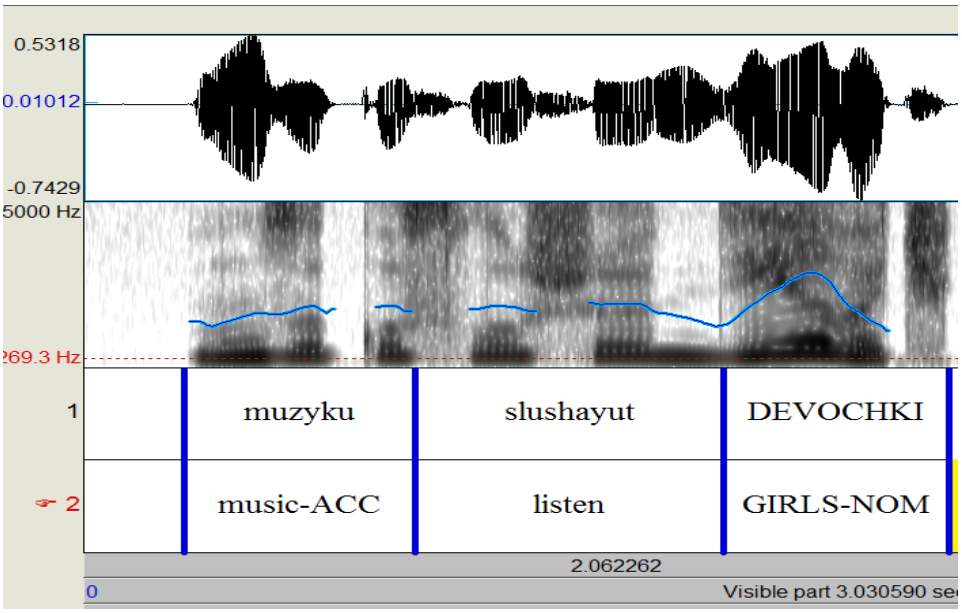


Figure 4.4. *Spectrogram of the stimulus in (4.1),non-canonical order, sentence-final subject NP is prominent*



(4.2) a. O[target noun, neutral prosody] VS c. O[target noun, prosodically prominent] VS

To verify that the differences between the prosodic conditions were robust, f0 (Hz) was sampled from all vowel and voiced consonant regions of subject and object nouns in the spliced stimuli. A sampling frequency of 100 samples per second was used. f0 values for the voiceless regions were supplied using cubic spline interpolation in Matlab. Tonal contours of subject and object nouns were time-normalized and reduced to 10 discrete time points using the normalization procedure reported in Wang, Jongman, and Sereno (2003).

Acoustic-prosodic measures of f0 range and maxima, mean intensity and vowel duration were taken from the stressed syllable of each subject and object word in the 32 test sentences. All measurements were extracted automatically in Praat (Boersma and Weenink 2013). Fundamental frequency and intensity values 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 Hz. Each prosodic measure entered a separate analysis of variance with predictor variables *Prosodic Prominence* (2 levels: target noun prominent or non-prominent), *Word Order* (2 levels: target noun in-situ or ex-situ).

For production data analyses, it is predicted that nouns which were purposefully uttered as prosodically prominent by the model speaker will show evidence for robust augmentation of all acoustic-prosodic parameters included into the analyses. Independently of the controlled prosodic manipulation and in line with prior research on Russian (Luchkina et al. 2015, Luchkina and Cole, forthcoming), it is predicted that the ex-situ nouns in the non-canonical OVS sentences will demonstrate evidence of partial acoustic-prosodic augmentation (evident, specifically, for fronted objects and post-posed subjects), possibly, in combination with partial acoustic-prosodic reduction (evident, specifically, for post-posed subjects).

Results

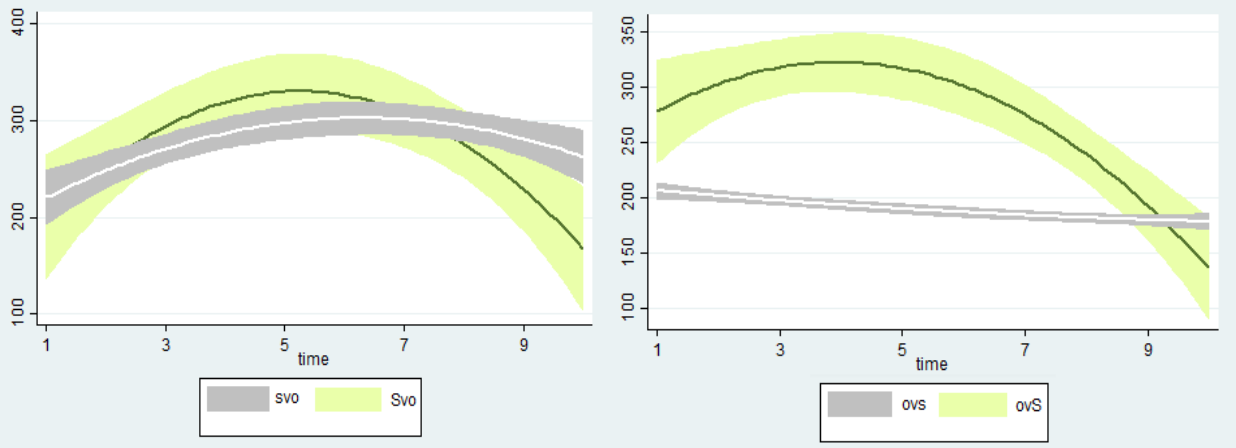
Figures 4.5. and 4.6. show two-way quadratic prediction plots with 95% confidence intervals for time-normalized f0 contours of subject (Figure 4.5.) and object (Figure 4.6.) target nouns in SVO and OVS stimuli, averaged across individual tokens. The left panel of each graph shows contours of sentence-initial target nouns, in-situ subjects and ex-situ fronted objects, and the right panel – of sentence-final target nouns, ex-situ post-posed subjects and in-situ objects. Visual examination of these contours reveals that the tonal characteristics of subject and object nouns in the experimental stimuli are affected by the linearization of sentence constituents, as well as by prosodic manipulation.

As predicted, keeping word order constant, prosodically prominent nouns in the model speaker's read production have greater f0 excursion (t=4.96, p<0.001), mean intensity (t=3.17, p<0.005),

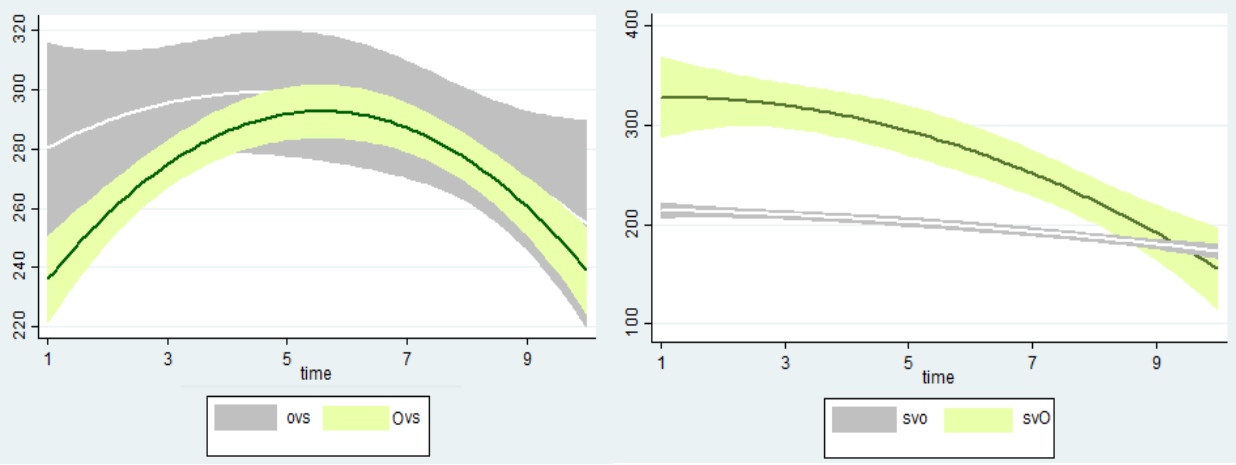
and duration (t=3.07, p<0.005). Regardless of prosodic manipulation, ex-situ nouns have consistently greater f0 excursion (t=2.8, p=0.005) and greater duration (t=3.1, p<.005). Additionally, acoustic-prosodic variability in the production data also depend on the grammatical function–animacy asymmetry in the subject and object nouns. Specifically, grammatical subjects, consistently animate, are more prodosically prominent than the inanimate grammatical objects and have greater f0 range (t=2.5, p<.05) and mean intensity (t=2.2, p<.05).

Figures 4.5. (upper panel) and 4.6. (bottom panel). *Stylized f0 contours of the target nouns in the experimental stimuli. Upper case letter in the word order abbreviation means that the target noun is prosodically prominent (e.g., subject in ovS).*

Subject nouns:



Object nouns:



To summarize, in the experimental sentences, prosodic prominence is most apparent for words which were read as prominent by the model speaker; however, more subtle prosodic augmentation is observed in animate subjects and ex-situ nouns in the non-canonical OVS order.

Probe recognition task

Method and participants

In an online probe recognition task, twenty-eight native Russian speakers heard SVO and OVS experimental sentences each followed by a lexical probe on a computer monitor. They were instructed to press YES if the probe had occurred in the test sentence, and to press NO otherwise. Response times (henceforth RTs, in ms) reflected the time interval between the auditory stimulus offset and the YES/NO button press.

For the task purposes, each experimental sentence was embedded into a carrier phrase such that the sentence-initial and sentence-final target nouns and the probes were separated with an equal amount of lexical material (in syllables). Each stimulus sentence was also preceded with a sentence-long vignette read by the same speaker, as shown in (1) above. Vignettes introduced the subject and the object nouns used in the SVO and OVS continuations, to minimize the effects of lexical frequency, word length, and cloze probability of the target noun on subsequent probe recognition. Critically, vignettes were designed such as to render the target noun in each experimental sentence contrastively prominent. Contrastive reading of the target noun was desirable for the following reasons. First, contrastively focused constituents in Russian are known to appear in- and ex-situ, with ex-situ occurrences localized to the sentence-initial, as well as sentence-final positions. Second, contrastive focus in Russian has been previously linked to prosodic augmentation of the focused constituent, evident from the expansion of f0 range and augmented duration (Botinis et al. 2005). Experimental manipulations introduced in this study are

highly consistent with the characteristics of contrastive focus in Russian. Contextually supporting constrastive reading of the target noun was used to ensure that the stimuli sounded maximally naturally.

Probes were 1-3 syllable lexical nouns; for test items, they always matched the target noun. The task also included 32 filler items created using the splicing procedure described in Section 2.1.1. Filler item contexts did not establish the subject or the object of the following sentence. Non-matching probes, semantically related to the subject or the object noun, were used in filler trials to encourage participants to pay attention during the task. Test and filler items were divided between three lists and pseudo-randomized. Stimuli were presented through headphones using E-prime2 software.

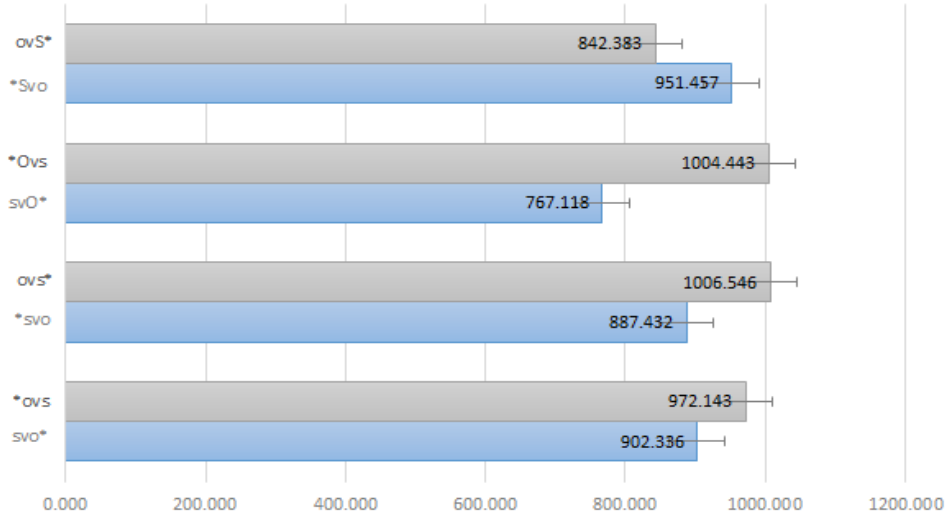
Native Russian speakers participated in the probe recognition task. Ten participants completed lists 1 and 2 and, 8 participants completed list 3. Responses from two participants were excluded from subsequent analyses due to systematically greater response times (approx. 20% of response times exceeded 3000 ms for 1 respondent) or reporting a different language as dominant (1 respondent). Data from 10 males and 16 females (29.6 y.o., *SD*=7.29) satisfied the inclusion criteria for subsequent analyses. All of these participants resided in the U.S. at the time of participation and were international students at a public university located in Illinois. The mean age of arrival to the US was 25.3 years (*SD*=5.99). All respondents reported being born and raised in Russia, in a Russian-speaking household. Two respondents had a bilingual parent who spoke Russian and Ukranian. One respondent had a bilingual parent who spoke Russian and Polish. Four respondents had bilingual parents who spoke Russian and Kazakh, and were themselves fluent speakers of Kazakh. All respondents reported Russian to be their native and dominant language.

Response times analyses

In the data from the remaining 26 participants, trials which resulted in response times which were greater than 3000 ms (<2% of all trials) and filler trials were not analyzed. Response times were modeled using a mixed effects linear regression. The model included fixed effects *Grammatical Function*, *Word Order* (target noun is in- or ex-situ), and *Prosodic Prominence* (target noun is/is not prominent) and a two-way interaction between *Word Order* and *Prosodic Prominence*. *Respondent* and *Test item* were introduced as random effects.

For the probe recognition task data analyses, consistent with earlier work on Russian (Sekerina 1999, Slioussar 2011a), it is predicted that response times should be smaller for canonically ordered SVO stimuli. Likewise, recognition of lexical probes should be facilitated whenever the target noun is prosodically prominent (see Fraundorf et al. 2010 for similar evidence for English). If prosodic augmentation and a change in word order complement each other, faster identification and recognition of lexical probes is expected when the target noun is ex-situ and prosodically prominent.

Figure 4.7. *Mean response times (ms) for SVO (blue) and OVS (grey) stimuli. Upper case letter in the word order abbreviation marks prosodically prominent word (e.g., Subject in ovS). Asterisk marks location of the target noun.*



Results

A mean accuracy rate of 98.5% indicates that participants paid attention. Response times (means and standard deviations) from 26 participants are summarized in Figure 4.7.

Consistent with the overall preference for the canonical SVO order in Russian, regression analysis revealed a significant main effect of *Word Order* (z=3.02, p<.005): Response times were smaller when the target noun occurred in-situ (in the canonical SVO order). Smaller RTs were also recorded whenever the probe matched the sentence subject as opposed to the sentence object (z=4.25, p<.001).

Word Order interacted with *Prosodic Prominence*. Smaller response times were obtained for probes matching (1) prosodically neutral subject or object noun in the S[neutral]VO[neutral] order (z=-4.96, p<.001); (2) prosodically prominent object in the S[neutral]VO[prominent] order (z=-5.11, p<.001), and (3) prosodically prominent subject in the O[neutral]VS[prominent] order (z=-3.47, p=.001). Significantly longer response times were obtained for probes matching prosodically prominent object in the O[prominent]VS order (z=4.24, p<.001). To summarize, for subject and object target nouns alike, an ex-situ position was associated with longer response times (see Figure 4.7). The only exception to this pattern were trials in which the probe matched a sentence-final subject noun which was also prosodically-prominent (OVS[prominent]).

Discussion

The focus of the present investigation is the acoustic-prosodic expression of ex-situ constituents in Russian, a free word order language, its status in discourse and relative contribution during auditory comprehension of a discourse or narrative. An experimental investigation involving production and perception data obtained from linguistically naïve native Russian speakers was carried out to determine if the special acoustic-prosodic expression of ex-situ words in Russian is

used during auditory sentence comprehension, possibly as a means of marking pragmatic focus and relative information prominence. Specifically, it was evaluated if auditory comprehension of non-canonical sentences featuring a prosodically-augmented ex-situ word leads the listener to identify that word faster as well as with greater accuracy.

Experimental materials used in this study included 16 canonical SVO and 16 non-canonical OVS sentences in Russian read by a female native speaker. Acoustic-prosodic correlates f0, duration, and mean intensity were examined in the acoustic-prosodic expression of in- vs. ex-situ subject and object nouns in the experimental stimuli. Analyses of the model speaker's read production data revealed that in the non-canonical OVS order, ex-situ subject and object nouns underwent partial prosodic augmentation and had greater duration and f0 range.

Since the ex-situ words in the experimental sentences were either sentence-initial or sentence-final, prosodic domain boundary strengthening may have had an effect on some or all of the acoustic-prosodic parameters of fronted objects and post-posed subjects. In this case, partial prosodic augmentation of the ex-situ nouns in the experimental stimuli may be regarded as an aftermath of deviant constituent linearization. A more dramatic acoustic-prosodic realization of ex-situ nouns may be further boosted by hyperarticulation effects due to added computational complexity associated with non-canonical orders in Russian. On a different account, augmented duration and expanded f0 range indicative of the ex-situ words in the non-canonical OVS order may be regarded as evidence for dual expression of their discourse status and relative information prominence.

To tap into the nature of the acoustic-prosodic effects observed in the read production data, an online probe recognition task exposed native Russian listeners to short discourse excerpts in which the experimental stimuli paired with short contexts rendering the target noun contrastively prominent. Analyses of the response times obtained in the probe recognition task revealed that

listeners demonstrated sensitivity to word order and acoustic-prosodic expression systematically varied in the test stimuli.

Consistent with the overall preference for the canonical SVO order in Russian, response times were shorter when the target noun occurred in canonically ordered SVO stimuli, regardless of its acoustic-prosodic characteristics. Response times were significantly longer when the target noun was presented in the non-canonical OVS order. This finding is consistent with the view that additional cognitive resources may be involved during processing of the non-SVO orders in Russian, possibly, due to the more complex syntactic derivation of such orders.

Notably, experimental investigations which examined the computational load associated with processing of non-canonically ordered sentences, using data from German, Japanese, Finnish, and Slavic languages, were not always able to support this view. To illustrate, in an online self-paced reading experiment, Featherston (1999) and Clahsen and Fetherston (1999) found that no additional processing costs were required for processing of the non-canonical orders in German, even though the stimuli used in those studies introduced temporary ambiguity due to homomorphous case marking. Similarly, Nakayama (1995) reported that in a probe recognition experiment involving non-canonical orders in Japanese, no additional time was required for processing non-canonically ordered sentences. Yamashita (1997) studied scrambling in Japanese and replicated Nakayama's finding in a series of self-paced reading experiments. An eye tracking study by Hyona and Hujanen (1997) offered evidence for equal processing costs associated with canonical and non-canonical word orders in Finnish: the authors reported that, unexpectedly, their respondents fixated on the ex-situ subjects for the *least* amount of time.

Turning now to work which presented evidence that greater processing costs are required for non-canonically ordered discourse segments, in a study carried out by Kaiser and Trueswell (2004),

Finnish respondents had to read canonical SVO and non-canonical OVS sentences preceded by two types of context, facilitative, in which subjects were discourse-new, or non-facilitative, in which subjects were discourse-given, rendering the change in word order unmotivated. Kaiser and Trueswell found that their participants' reading times decreased when the ex-situ subject noun in the OVS stimuli was discourse-new, while the object noun was discourse-given. The authors (2004) proposed that processing of non-canonical orders required additional memory resources: having encountered an object pre-verbally, speakers of an SVO language such as Finnish, have to actively maintain an awareness of a delayed subject until its surface representation is reached later in the sentence.

Sekerina (1999) and Slioussar (2011a) looked at the processing of written context appropriate sentences with canonical and non-canonical orders in Russian and found that significantly greater reading times were associated with non-canonically-ordered sentences, such as OSV (Sekerina 1999, Slioussar 2011a), INDIRECTOSVODIRECT, and INDIRECTODIRECTOVS (Slioussar 2011a). In line with this prior work on Russian, this study documents a reliable response time advantage associated with canonically-ordered sentences compared to the scrambled OVS order, regardless of the prosodic manipulation discussed next.

Recall that due to controlled prosodic manipulation, 50% of the target nouns in the experimental stimuli had a comprehensively augmented acoustic-prosodic expression. Results of the probe recognition task revealed that acoustic-prosodic augmentation of a sentence-final noun, subject or object, was associated with significantly smaller response times, in the canonical S[neutral]VO[prominent] and the scrambled OVS order (see Figure 4.7). This result is not surprising, since in many head-initial languages, including Russian, the sentence-final position presents the preferred location of the major phrasal prominence or nuclear pitch accent and acts as

a natural prominence landing cite. Smaller response times obtained for S[neutral]VO[prominent] and O[neutral]VS[prominent] stimuli suggest that preservation of the underlying prosodic structure at the phrasal level, just like the preservation of the canonical constituent ordering in Russian, facilitated faster response times.

Listeners selectively benefitted from the cross-application of word order and acoustic-prosodic variability. To illustrate, response times were significantly smaller when the sentence-final subject noun was prosodically augmented in the O[neutral]VS[prominent] order, pointing to an overall facilitative effect of prosodic accenting on the subsequent recognition of the target noun. Likewise, in the canonical SVO order, making the utterance-final object prosodically prominent was conducive to faster probe recognition. However, prosodic prominence aligned with the sentence-initial object in the O[prominent]VS[neutral] did not result in faster probe recognition, but was instead associated with significantly *longer* response times.

One possible explanation behind the divergent effects of prosodic prominence in the non-canonical OVS order is that sentence-initial inanimate objects in the O[prominent]VS[neutral] stimuli were interpreted as (contrastive) sentence topics, for which accenting could be regarded as suboptimal. Alternatively, the sentence-initial position may present a suboptimal prominence landing cite in Russian, compared to the sentence-final position which, by default, is nuclear pitch-accented. This latter interpretation is supported with the observation that numerically, the S[prominent]VO[neutral] order was also associated with longer response times as well, although this trend did not reach significance.

Results of the probe recognition task reveal that a change in the surface constituent order in Russian does not preclude listeners from deploying concurrent acoustic-prosodic variability observed in relation to discourse status and relative information prominence of the ex-situ constituent.

However, under a non-canonical word order, the facilitative effect of prosodic cues on the recognition of the ex-situ word was observed only when the latter moved to a natural prominence-landing site, such as the utterance-final position in Russian. These results are consistent with the proposal that acoustic-prosodic expression of ex-situ fronted objects in the non-canonical OVS order may be attributable to domain-initial prosodic strengthening, per Cutler, Dahan, Van Donselaar (1997), or reflected a more effortful production of the non-canonical orders. Further experimental work is needed to tease these accounts apart for Russian.

Conclusion

In a free word order language, information of relatively greater perceived prominence may be manifest by prosodic means, reflective of the underlying phonological structure of a language, as well as by linearization means, reflective of its underlying syntactic structure. While in some languages, only one prominence marking mechanism, acoustic-prosodic or structural may apply to the exclusion of the other (Swerts et al. 2002, Donati and Nespor 2003), in others, both these mechanisms may not only 'co-exist', but also cross-apply (Skopeteas et al., 2010, Arvaniti & Adamou, 2011, Botinis et al. 2005). Not surprisingly, when prosodic and structural cues to prominence are simultaneously available in a language, they appear highly interrelated.

Analyses of read production data presented in this study and in the prior work successfully capture the close interrelatedness of prosodic and structural variability in read discourse in Russian, a highly free word order language known for discourse-conditioned word order variability. In line with the previous work on Russian (Botinis et al. 2005, Luchkina and Cole, forthcoming), this study reports that altering the order of sentence constituents partially augments the acoustic-prosodic realization of the ex-situ word. While the cross-application of acoustic-prosodic and word order variability is attended to by the listener, it only facilitates subsequent recognition of the ex-

situ word when the latter is aligned with a natural prominence-landing site, such as the utterance-final position in Russian. It appears that departure from the preferred (canonical) constituent order, while computationally costly, may be implemented to align the ex-situ word with the designated location of the major phrasal prominence, where the special discourse status and the relative information prominence of the ex-situ word are clearly manifest.

REFERENCES

Antonyuk-Yudina, S., & R. Mykhaylyk. (2013): Prosodic Effects in Word Order: Evidence from Ukrainian. In *Proceedings of North East Linguistic Society 40*, MIT, Cambridge, MA.

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.

Arvaniti, A., & Adamou, E. (2011). Focus expression in Romani. In *Proceedings of the 28th West Coast conference on Formal Linguistics*, Cascadilla Proceedings Project, Somerville, MA.

Ayers, G. (1996). *Nuclear accent types and prominence: Some psycholinguistics experiments.* Unpublished doctoral dissertation. The Ohio State University, Columbus, OH.

Aylett, M., & Turk, A. (2004). The smooth signal redundancy hypothesis: A functional explanation for relationships between redundancy, prosodic prominence, and duration in spontaneous speech. *Language and Speech* 47: 31-56.

Bailyn, J. (2002). Inversion, dislocation and optionality in Russian. In Zybatow, G. et al. (eds.) *Current Issues in Formal Slavic Linguistics (FDSL 3)*. Frankfurt am Main, Peter Lang: 280-293.

Bailyn, J. (2003). A (purely) derivational account of Russian scrambling. In Browne, W., Kim, J-Y., Partee, B., & Rothstein, S. (eds.) *Formal Approaches to Slavic Linguistics (FASL) 11*, Ann Arbor, MI: Michigan Slavic Publications: 41-62.

Bailyn, J. (2004). Generalized inversion. *Natural Language and Linguistic Theory* 22: 1-49.

Baltazani, M. (2003). Pragmatics, intonation, and word order in Greek. In *Proceedings of Interfaces prosodiques IP 2003*: 14-18.

Baltazani, M., & Jun, S-A. (1999). Focus and topic intonation in Greek. In *Proceedings of the XIVth International Congress of Phonetic Sciences 2*: 1305-1308.

Bard, E.G., Anderson, A.H., Sotillo, C., Aylett, M., Doherty-Sneddon, G., & Newlands, A. (2000). Controlling the Intelligibility of Referring Expressions in Dialogue. *Journal of Memory and Language* 42-1: 1-22.

Bard, E.G., & Aylett, M.P. (1999). The Dissociation of Deaccenting, Givenness and Syntactic Role in Spontaneous Speech. In *Proceedings of ICPhS-99*, San Francisco, CA.

Baumann, S. (2006). Information Structure and Prosody: Linguistic Categories for Spoken Language Annotation. In: Sudhoff, S., Lenertová, D., Meyer, R., Pappert, S., Augurzky, P., Mleinek, I., Richter, N., & Schließer, J. (eds.), *Methods in Empirical Prosody Research*. Berlin, New York: De Gruyter: 153-180.

Baumann, S., Riester, A. (2012). Referential and Lexical Givenness: Semantic, Prosodic and Cognitive Aspects. In Elordieta, G., & Prieto, P. (eds.), *Prosody and Meaning*. Mouton De Gruyter, Berlin, New York.

Baumann, S., Riester, A. (2013). Coreference, lexical givenness and prosody in German. *Lingua*

136: 16–37.

Baumann, S., & Riester, A. (2012). Referential and Lexical Givenness: Semantic, Prosodic and Cognitive Aspects. In Elordieta, G., & Prieto, P. (eds.) *Prosody and Meaning*. Mouton De Gruyter, Berlin, New York.

Birner, B. (1994): Information status and word order: An analysis of English inversion. *Language*, 70: 233-259.

Bishop, J. (2012). Information structural expectations in the perception of prosodic prominence. In G. Elordieta; P. Prieto (eds.) Prosody and meaning (Trends in Linguistics), Mouton de Gruyter, Berlin, New York.

Bivon, R. (1971). Element Order. *Studies in the Modern Russian Language*, 7. Cambridge University Press, Cambridge.

Bock, J. K., Warren, R. K. (1985). Conceptual accessibility and syntactic structure in sentence formulation. *Cognition* 21: 47-67.

Boersma, P., & Weenink, D. (2016). Praat: doing phonetics by computer [Computer program]. Version 6.0.14, retrieved on 11 February 2016 from http://www.praat.org/

Bolinger, D. (1986). *Intonation and its Parts: Melody in Spoken English*. Stanford, CA: Stanford University Press.

Bornkessel-Schlesewsky, I., & Schlesewsky, M. (2009). The role of prominence information in the real-time comprehension of transitive constructions: A cross-linguistic approach. *Language and Linguistic Compass* 3/1: 19-58.

Botinis, A., Themistocleous, C., Kostopoulos, Y., & Nikolaenkova, O. (2005). Syntactic and tonal correlates of focus in Greek and Russian. In *Proceedings Fonetik 2005, the XVIIIth Swedish Phonetics Conference*. Goteborg, Sweden.

Branigan, H.P., Pickering, M.J., & Tanaka, M. (2008). Contributions of animacy to grammatical function assignment and word order during production. *Lingua*, 118: 172-189.

Branigan, H.P., & Feleki, E. (1999). Conceptual accessibility and serial order in Greek language production. In: *Proceedings of the 21st Conference of the Cognitive Science Society, Vancouver.*

Breen, M., Fedorenko, E., Wagner, M., & Gibson, E. (2010). Acoustic correlates of information structure. *Language and Cognitive Processes* 25: 1044-1098.

Bresnan, J., Dingare, S., & Manning, C. D. (2001). Soft constraints mirror hard constraints: Voice and person in English and Lummi. In *Proceedings of the LFG '01 Conference*. CSLI Publications.

Brun, D. (2001). Information structure and the status of NP in Russian. *Theoretical Linguistics* 27: 109-135.

Bryzgunova, E. A. (1980). *Intonacija*. In: Švedova, N. J. (Ed.), Russkaja grammatika. Tom 1. Nauka, Moskva.

Bush, L., Hess, U., & Wolford, G. (1993). Transformations for within-subject designs: A Monte Carlo Investigation. *Psychological Bulletin* 113: 566-579.

Calhoun, S. (2010). The Centrality of Metrical Structure in Signaling Information Structure: A Probabilistic Perspective, *Language*, 86(1): 1-42.

Calhoun, S. (2015). The interaction of prosody and syntax in Samoan focus marking. *Lingua*, special issue: prosody and information status in typological perspective, edited by Frank Kügler & Stefan Baumann.

Calhoun, S., M., Nissim, M., Steedman, & Brenier, J. (2005). A Framework for Annotating Information Structure in Discourse. In *ACL 2005 Workshop: Frontiers in Corpus Annotation II*.

Cameron, C. A., & Trivedi, P. K. (2013). *Regression Analysis of Count Data*. Econometric Society Monograph 53 Cambridge University Press, Cambridge

Chafe, W. L. (1994). *Discourse, consciousness and time*. The University of Chicago Press, Chicago/London.

Chafe, W.L. (1976). Givenness, contrastiveness, definiteness, subjects, topics, and point of view. In Li, C. (ed.), *Subject and Topic*. Academic Press, New York.

Cho, T. (2016). Prosodic boundary strengthening in the phonetics-prosody interface. *Language and Linguistics Compass*, 10(3): 120-141.

Clark, E., & Clark, H. (1978). Universals, relativity, and language processing. In J. H. Greenberg (Ed.), *Universals of human language*. Stanford, CA: Stanford University Press.

Clark, H. H., & Haviland, S. E. (1977). Comprehension and the given-new contract. In R. O. Freedle (ed.), *Discourse production and comprehension*. Hillsdale, NJ: Erlbaum: 1-40.

Clifton, C., & Frazier L. (2004). Should given information come before new? Yes and no. *Memory & Cognition* 32: 886-895.

Cole, J., Mo, Y., & Hasegawa-Johnson, M. (2011). Signal-based and expectation-based factors in the perception of prosodic prominence. *Laboratory Phonology* 1: 425-452.

Comrie, B. (1989). *Linguistic universals in language typology* (2nd ed.) Blackwell, Oxford, UK.

Cruttenden, A. (2006). The deaccenting of given information: a cognitive universal? In Bernini, G., Schwartz, M. (eds.), *Pragmatic organization of discourse in the languages of Europe*. Mouton de Gruyter, The Huage: 311-356.

Cutler, A., Dahan, D., & Van Donselaar, W. (1997). Prosody in the comprehension of spoken language: A literature review. *Language and Speech*, 40:141-201.

de Ruiter, L. E. (2015). Information status marking in spontaneous vs. read speech in story-telling tasks – Evidence from intonation analysis using GToBI. *Journal of Phonetics*, 48: 29-44.

de Swart, P. (2007). *Cross-linguistic variation in object marking*. LOT, Utrecht, The Netherlands.

Dilley, L. (2005). *The phonetics and phonology of tonal systems*. Ph.D. Dissertation, MIT.

Donati, C., & Nespor, M. (2003). From focus to syntax. *Lingua*, 113.11: 1119-1142.

Dragoy, O., & Bastiaanse, R. (2009). Verb production and word order in Russian agrammatic speakers. *Aphasiology,* 24: 28–55.

Erdocia, K., Laka, I., Mestres-Missé, A, & Rodriguez-Fornells, A. (2009). Syntactic complexity and ambiguity resolution in a free word order language: behavioural and electrophysiological evidences from Basque. *Brain and Language*, 109 (1): 1-17.

Ertschik-Shir, N. (2007). *Information structure: the syntax-discourse interface*. Oxford: Oxford University Press.

Ferreira, F. (2003). The misinterpretation of noncanonical sentences. *Cognitive Psychology* 47: 164–203.

Féry, C. (2013). Focus as prosodic alignment. *Natural Language and Linguistic Theory* 31(3): 683–734.

Féry, C., & Krifka, M. (2008). 'Information structure: Notional distinctions, ways of expression.' In van Sterkenburg, Piet (ed.), *Unity and diversity of languages*. Amsterdam: John Benjamins.

Fisher, C., & Tokura, H. (1995). The given/new contract in speech to infants. *Journal of Memory and Language*, 34: 287-310.

Fowler, C. A., & Housum, J. (1987). Talker's signaling of 'new' and 'old' words in speech and listeners' perception and use of the distinction. *Journal of Memory and Language*, 26(5): 489-504.

Fraundorf, S.H., Watson, D.G., & Benjamin, A.S. (2010). Recognition Memory reveals just how CONTRASTIVE contrastive accenting really is. *Journal of Memory and Language*, 63: 367-386.

Frazier, L., & Clifton, C. (1996). *Construal*. MIT Press, Massachusetts.

Frisch, S., & Schlesewsky, M. (2001). The N400 indicates problems of thematic hierarchizing. *Neuroreport* 12: 3391-3394.

Frota, S. (2002). The prosody of focus: a case-study with cross-linguistic implications. In *Proceedings of Speech Prosody 2002*: 315-318.

Genzel, S. Ishihara, B., & Surányi, B. (2015). The prosodic expression of focus, contrast and givenness: A production study of Hungarian. *Lingua* 165: 183-204.

Genzel, S., & Kügler, F. (2010). The prosodic expression of contrast in Hindi. In *Proceedings of Speech Prosody 2010.*

Gordon, P. C., Hendrick, R., & Johnson, M. (2001). Memory interference during language processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 27(6): 1411–1423.

Gow, D. W., Melvold, J., & Manuel, S. (1996). How word onsets drive lexical access and segmentation: evidence from acoustics, phonology and processing. In *Proceedings of ICSLP* (CD ROM). Philadelphia, USA.

Gussenhoven, C. (2007). Types of Focus in English. In Chungmin Lee, Matthew Gordon, and Daniel Büring (eds.) *Topic and Focus: Cross-linguistic Perspectives on Meaning and Intonation*. Heidelberg, New York, London: Springer: 83-100.

Halliday, M. (1967). *Intonation and grammar in British English*. The Hague: Mouton.

Hanssen, J., Peters, J., & Gussenhoven, C. (2008). Prosodic effects of focus in Dutch declaratives. In *Proceedings of Speech Prosody 2008*. Campinas, Brazil: 609-612.

Haviland, S., & Clark, H. (1974). What's new? Acquiring new information as a process in comprehension. *Journal of Verbal Learning and Verbal Behavior*, 13, 512-521.

Hermes D.J., & Rump H.H. (1994). Perception of prominence in speech intonation induced by rising and falling pitch movements. *Journal of the Acoustical Society of America* 96: 83-92.

Hinterhölzl, R. (2009). Information structure and unmarked word order in (Older) Germanic. In Zimmermann, M.; C. Féry (eds.) *Information Structure: Theoretical, Typological, and Experimental Perspectives*. Oxford University Press, Oxford.

Hoeks, J. C. J., Stowe, L. A., & Doedens, G. (2004). Seeing words in context: the interaction of lexical and sentence level information during reading. *Cognitive Brain Research* 19: 59–73.

Hróarsdóttir, T. (2009). Information structure and OV order. In Zimmermann, M. and C. Féry (eds.) *Information Structure: Theoretical, Typological, and Experimental Perspectives*. Oxford University Press, Oxford.

Ionin, T. (2001). The One Girl who was Kissed by Every Boy: Scope, Scrambling and Discourse Function in Russian, in Marjo van Koppen, Joanna Sio, and Mark de Vos (eds.), *Proceedings of Console X. Student Organization of Linguistics in Europe*. Leiden University. Leiden, The Netherlands, 79-94.

Ionin, T., & Luchkina, T. Focus on Russian Scope: An Experimental Investigation of the Relationship between Quantifier Scope, Prosody, and Information Structure. Manuscript under review.

Işsever, S. (2003). Information structure in Turkish: the word order-prosody interface. *Lingua*, 113: 1025-1053.

Jackendoff, R. (1972). *Semantic Interpretation in Generative Grammar*. Cambridge: MIT Press.

Jasinskaja, K. (2013). Information Structure in Slavic. In C. Féry, & Ishihara, S. (eds.) *Handbook of Information Structure*. Oxford University Press, Oxford.

Jun, S.-A. (2014). *Prosodic Typology II: the phonology of intonation and phrasing*. Oxford: Oxford University Press.

Kaland, C.C.L., Krahmer, E.J., & Swerts, M.G.J. (2011). Contrastive intonation: Speaker- or listener-driven? In W.-S. Lee & E. Zee (eds.), *Proceedings of the 17th International Congress of Phonetic Sciences*, 1006-1009.

Kallestinova, E. (2007). *Aspects of word order in Russian*. Unpublished doctoral dissertation, Iowa

University.

Katz, J., & Selkirk, E. (2011). Contrastive focus vs. discourse-new: evidence from phonetic prominence in English. *Language*, 87:4: 771-816.

Keating, P., Cho, T., Fougeron, C., & Hsu, C. (1999). Domain-initial strengthening in four languages. *UCLA Working Papers in Phonetics* 97: 139-151.

King, T. (1995). *Configuring Topic and Focus in Russian*. CSLI Publications: Center for the Study of Language and Information, Stanford, California.

Kiss, E. K. (1998). Identificational focus versus information focus. *Language* 74(2): 245–273.

Kochanski, G., Grabe, E., Coleman, J., & Rosner, B. (2005). Loudness predicts prominence: Fundamental frequency lends little. *Journal of the Acoustic Society of America* 118: 1038-1054.

Krifka, M. (2007). Basic notions of information structure. In Féry, C., Krifka, M. (eds.), *Interdisciplinary Studies of Information Structure* 6, Potsdam.

Krylova, O., & Khavronina, S. (1988). *Word Order in Russian*. Moscow: Russky Yazyk.

Kuperberg, G. R., Sitnikova, T., Caplan, D., & Holcomb, D. (2003). Electrophysiological distinctions in processing conceptual relationships within simple sentences. *Cognitive Brain Research* 17: 117–129.

Kuperberg, G. R., Kreher, D. A., Sitnikova, T., Caplan, D. N., & Holcomb, P. J. (2007). The role of animacy and thematic relationships in processing active English sentence: evidence from event-related potentials. *Brain and Language* 100: 223–37.

Ladd, R. D. (2008). *Intonational Phonology*. Cambridge University Press, Cambridge, MA.

Lambrecht, K. (1994). *Information Structure and Sentence Form*. Cambridge University Press, Cambridge, MA.

Lobanova, A. (2011). The role of prominence scales for the disambiguation of grammatical functions in Russian. *Russian Linguistics*, 35-1: 125-142.

Luchkina, T., & Cole, J. (forthcoming). Structural and Referent-based Effects on Prosodic Expression in Russian. *Phonetica*.

Luchkina, T., Puri, V., Jyothi, P., & Cole, J. (2015). Prosodic and structural correlates of perceived prominence in Russian and Hindi. In *Proceedings of International Congress of Phonetic Sciences.*

Luchkina, T., Cole, J. (2014). Structural and Prosodic Correlates of Prominence in Free Word Order Language Discourse. In *Proceedings of the 2014 meeting of the Speech Prosody Conference (SP 7)*, Dublin, Ireland.

Luchkina, T., Puri, V., Jyothi, P., & Cole, J. (2015). Prosodic and structural correlates of perceived prominence in Russian and Hindi. In *Proceedings of International Congress of Phonetic Sciences 2015*.

Luchkina, T., & Cole, J. Prosodic and Structural Correlates of Perceived Prominence in a Free

Word Order Language. Manuscript in preparation.

MacWhinney, B., & Bates, E. (1978). Sentential devices for conveying givenness and newness. *Journal of Verbal Learning and Verbal Behavior*, 17, 539-558.

Mak, W. M., Vonk, W., & Schriefers, H. (2002). The influence of animacy on relative clause processing. *Journal of Memory and Language*, 47: 50–68.

Makarova, V. (2007). The effect of pitch peak alignment contrasts on sentence type identification in Russian. *Language and Speech*, 50(3), 385-422.

Marslen-Wilson W. D., Levy, E., & Tyler, L. K. (1982). Producing interpretable discourse: The establishment and maintenance of reference. In Jarvella, R.J., & Klein, W. (eds.), *Speech, Place, & Action.* London: Wiley.

McDonalnd, J. L., Bock, K., & Kelly, M. H. (1993). Word and world order: semantic, phonological, and metrical determinants of serial position. *Cognitive Psychology*, 25: 188-230.

McKoon, G., Ratcliff, R., Ward, G., & Sproat, R. (1993). Syntactic prominence effects on discourse processing. *Journal of Memory and Language*, 32, 593-607.

Meyer, R., & Mleinek, I. (2006). How prosody signals force and focus — A study of pitch accents in Russian yes–no questions. *Journal of Pragmatics*, 38: 1615-1635.

Mitchell, M. N. (2012). *Interpreting and Visualizing Regression Models Using Stata*. Stata Press, College Station, Texas.

Miyamoto, E. T., & Takahashi, S. (2004). Filler-gap dependencies in the processing of scrambling in Japanese. *Language and Linguistics*, 5, 153-166.

Mo, Y. (2008). Duration and intensity as perceptual cues for naïve listeners' prominence and boundary perception. In *Proceedings of the 4th Speech Prosody*, Campinas, Brazil.

Mo, Y., Cole, J., & Lee, E. (2008). Native listeners? Prominence and boundary perception. In *Proceedings of Speech Prosody 2008*.

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.

Myers, R. (1990). *Classical and Modern Regression with Applications*. Duxbury Press, Boston, MA.

Nakayama, M. (1995). Scrambling and probe recognition. In R. Mazuka & N. Nagai (eds), *Japanese Sentence Processing*. Lawrence Erlbaum Associate, Hillsdale, NJ. 257-273.

Neeleman, A., & Reinhart, T. (1998). Scrambling and the PF-interface. In Butt, M., & Geuder, W. (eds.) *The projection of arguments*: 309–353. CSLI.

Neeleman, A., & Titov, E. (2009). Focus, contrast, and stress in Russian. *Linguistic Inquiry* 40: 514–524.

Patil, U., Kentner, G., Gollrad, A., Kügler, F., Féry, C., & Vasishth, S. (2008). Focus, word order, and intonation in Hindi. *Journal of South Asian Linguistics*, 1: 53-70.

Pierrehumbert, J., & Hirschberg, J. (1990). The Meaning of Intonational Contours in the Interpretation of Discourse. In P. Cohen, J. Morgan, & M. Pollack, (eds). *Intentions in Communication*, MIT Press, Cambridge MA: 271-311.

Pierrehumbert, J., & Talkin, D. (1992). Lenition of /h/ and glottal stop. In: Gerard J. Doherty and D. Robert Ladd (eds.), Gesture, Segment, Prosody. In *Papers in Laboratory Phonology 2*. Cambridge: Cambridge University Press: 90-117.

Pluymaekers, M., Ernestus, M., & Baayen, R. (2005). Articulatory planning is continuous and sensitive to informational redundancy. *Phonetica* 62, 146–159.

Prat-Sala, M. (1997). *The production of different word orders: a psycholinguistic and developmental approach.* Ph.D. Dissertation. University of Edinburgh, Edinburgh.

Prat-Sala, M., & Branigan, H. P. (2000). Discourse constraints on syntactic processing in language production: A cross-linguistic study in English and Spanish. *Journal of Memory and Language*, 42: 168–82.

Prince, E. (1981). Toward a taxonomy of given/new information. In Cole, P. (ed). *Radical Pragmatics.* New York: Academic Press. 223-254.

Raftery, A. E. (1995). Bayesian model selection in social research. In P. Marsden (ed). *Sociological Methodology*. Washington, DC: The American Sociological Association: 111–163.

Riester, A., & Piontek, J. (2015). Anarchy in the NP. When new nouns get deaccented and given nouns don't. *Lingua*, 165(B): 230-253.

Rietveld, A.C.M., & Gussenhoven, C. (1985). On the relation between pitch excursion size and pitch prominence. *Journal of Phonetics*, 13: 299-308.

Ross, J. R. (1967). *Constraints on Variables in Syntax*. Doctoral dissertation, MIT. [Published in 1986 as Infinite Syntax. Norwood, N.J.: ABLEX.] Excerpted in Gilbert Harman (ed.), *On Noam Chomsky - Critical Essays*. Anchor Press/Doubleday, 1974: 165--200.

Sekerina, I. A. (1999). The Scrambling Complexity Hypothesis and Processing of Split Scrambling Constructions in Russian. *Journal of Slavic Linguistics* 7.2: 218-265.

Sekerina, I. A. (2003). Scrambling processing: Dependencies, complexity, and constraints. In S. Karimi (ed.), Word order and scrambling Blackwell, UK. 301–324.

Sekerina, I. A., & Trueswell, J. C. (2012). Interactive processing of contrastive expressions by Russian children. *First Language*, 32(1/2): 63–87.

Selkirk, E. (2011). The Syntax-Phonology Interface. In J. Goldsmith, J; Riggle, J.; Yu, A. (eds). *The Handbook of Phonological Theory*, 2nd edition. Blackwell Publishing, Oxford.

Siewierska, A. (1993). Syntactic weight vs. information structure and word order variation in Polish. *Journal of Linguistics*, 29: 233-265.

Sirotinina, O. B. (1965/2003). *Porjadok slov v russkom jazyke* ('Word order in Russian') 2nd ed. Editorial URSS, Moscow.

Sityaev, D. (2000). The relationship between accentuation and information status of discourse referents: A corpus-based study. In *UCL Working Papers in Linguistics* 12: 285-304.

Skopeteas, S., & Fanselow, G. (2008). Focus types and argument asymmetries: A cross-linguistic tudy in language production. In Carsten, B. (ed.) *Contrastive information structure analysis*, Amsterdam: Benjamins.

Skopeteas, S., & Fanselow, G. (2009). Effects of givenness and constraints on free word order language. In Zimmermann, M., & C. Féry (eds.) *Information Structure: Theoretical, Typological, and Experimental Perspectives.* Oxford University Press, Oxford.

Skopeteas, S., & Fanselow, G. (2010). Focus in Georgian and the expression of contrast. *Lingua*, 120: 1370-1391.

Skopeteas, S., Féry, C., & Asatiani, R. (2009). Word order and intonation in Georgian. *Lingua* 119: 102-127.

Slioussar, N. (2010). Russian data call for relational Information Structure notions. In: Zybatow G. et al. (eds.). Formal Studies in Slavic Linguistics. In *Proceedings of Formal Description of Slavic Languages 7*: 329-344. Peter Lang, Frankfurt am Main.

Slioussar, N. (2011a). Processing of a free word order language: The Role of Syntax and Context. *Journal of Psycholinguistic Research*, 40: 291-306.

Slioussar, N. (2011b). *Grammar and Information Structure: A novel view based on Russian data*. Utrecht institute of Linguistics OTS and St. Petersburg State University.

Švedova, N. Yu. (Ed.) (1982). *Russkaja Grammatika. II.* (in Russian. Russian Grammar. II). The USSR Academy of Sciences. Moscow: Nauka.

Svetozavora, N. (1998). Russian Intonation. In Hirst, D., & Di Cristo, A. (ed.) *Intonation systems: a survey of twenty languages*. Cambridge University Press, Cambridge, MA.

Swerts, M. (2007). Contrast and accent in Dutch and Romanian. *Journal of Phonetics*, 35: 380-397.

Swerts, M., Krahmer, E.J., & Avesani, C. (2002). Prosodic marking of information status in Dutch and Italian: A comparative analysis. *Journal of Phonetics* 30: 629-654.

Swerts, M., Strangert, E., & Heldner, M. (1996). F0 declination in spontaneous and read-aloud speech. In *Proceedings of the 4th international conference on spoken language processing (ICSLP)*. Philadelphia, PA, 1501-1504.

Szendrői, K. (2001). *Focus and the syntax-phonology interface*. Doctoral dissertation, University College London, London.

Szendrői, K. (2003). A stress-based approach to the syntax of Hungarian focus. *The Linguistic Review*, 20(1): 37-78.

Tanaka, M., Branigan, H.P., & Pickering, M.J. (2005). The role of animacy in Japanese sentence production. Paper Presented at the CUNY Sentence Processing Conference, Tucson, AZ

Terken, J., & Hermes, D. (2000). The perception of prosodic prominence. In Horne, M. (ed.), *Prosody: Theory and experiment*. Dordrecht: Kluwer Academic Publishers: 89–127.

Titov, E. (2012). *Information structure of argument order alternations*. University College London PhD thesis.

Traunmüller, H., & Eriksson, A. (1995). The frequency range of the voice fundamental in the speech of male and female adults. Unpublished Manuscript (retrieved from http://www.ling.su.se/staff/hartmut/aktupub.htm).

Traxler, M., Morris, M. J., Robin, K., & Seely, R. E. (2002). Processing Subject and Object Relative Clauses: Evidence from Eye Movements. *Journal of Memory and Language*, 47.69-90.

Vainio, M., & Järvikivi, J. (2006). Tonal features, intensity, and word order in the perception of prominence. *Journal of Phonetics*, 34: 319-342.

Vainio, M., & Järvikivi, J. (2007). Focus in production: Tonal shape, intensity, and word order. *Journal of the Acoustical Society of America*, 121(2): EL55-61.

van Gelderen, E. (2003). Asp(ect) in English modal complement. *Studia Linguistica*, 57.1

van Nice, K.Y., & Dietrich, R. (2003). Task sensitivity of animacy effects: evidence from German picture descriptions. *Linguistics*, 41: 825–849.

van Valin, R. D. Jr., & LaPolla, R. J. (1997). *Syntax: structure, meaning and function*. New York: Cambridge University Press.

Verhoeven, E. (2009). Subjects, agents, experiencers, and animates in competition: Modern Greek argument order. *Linguistische Berichte 219*: 355-376.

Verhoeven, E. (2014). Thematic prominence and animacy asymmetries. Evidence from a cross-linguistic production study. *Lingua*, 143: 129-161.

Wang, Y., Jongman, A., & Sereno, J. A. (2003). Acoustic and perceptual evaluation of Mandarin tone productions before and after perceptual training. *The Journal of the Acoustical Society of America*, 113.2: 1033-1043.

Watson, D. G. (2010). The many roads to prominence: Understanding emphasis in conversation. In Ross, B. (ed.). *The Psychology of Learning and Motivation*, Elsevier: 163-183.

Watson, D., Arnold, J.A., & Tanenhaus, M.K. (2008). Tic Tac TOE: Effects of predictability and importance on acoustic prominence in language production. *Cognition* 106: 1548-1557.

Weckerly, J., & Kutas, M. (1999). An electrophysiological analysis of animacy effects in the processing of object relative sentences. *Psychophysiology*, 36: 559–70.

Williams, R. (2015). Scalar Measures of Fit: Pseudo R2 and Information Measures (AIC & BIC) University of Notre Dame. Accessed from https://www3.nd.edu/~rwilliam/stats3/L05.pdf

Xu, Y. (1999). Effects of tone on the formation and alignment of f0 contours. *Journal of Phonetics*, 27: 22-106.

Xu, Y., & Xu, C. (2005). Phonetic realization of focus in English declarative intonation. *Journal of Phonetics*, 33: 159-197.

Yamashita, H. (1997): The effects of word-order and case-marking information on the processing of Japanese. *Journal of Psycholinguistic Research.* 26: 163–88.

Yokoyama, O. (1986). *Discourse and word order*. John Benjamins, Amsterdam.

Zubizarreta, M. L. (1998). *Prosody, Focus, and Word Order*. Cambridge, Mass.: MIT Press.

Zybatow G. Formal Studies in Slavic Linguistics. In *Proceedings of Formal Description of Slavic Languages 7*: 329-344. Peter Lang, Frankfurt am Main.

APPENDIX A: *Predicted marginal means for parameters f0 range (panel A), duration (panel B) and mean intensity (panel C) across levels of AGRC*

Panel A: f0 range (ST)

speaker ID	object, animate	object, inanimate	subject, animate
1	3.52	1.78	4.33
	(.68)	(.63)	(.46)
2	2.15	2.78	3.02
	(.37)	(.28)	(.3)
3	2.22	2.37	3
	(.37)	(.27)	(.3)
4	1.7	1.84	2.39
	(.42)	(.34)	(.36)
5	1.75	2.33	2.43
	(.37)	(.28)	(.3)
6	3.95	3.52	4.38
	(.41)	(.3)	(.36)
7	2.32	2.82	3.15
	(.44)	(.32)	(.32)
8	3.44	2.43	3.35
	(.37)	(.28)	(.29)
9	2.47	2.02	2.7
	(.29)	(.21)	(.24)
10	2.52	2.5	3.2
	(.37)	(.29)	(.31)
11	1.58	1.85	2.4
	(.39)	(.28)	(.31)
12	1.4	1.76	2.23
	(.37)	(.28)	(.3)
13	3.53	2.14	2.52
	(.38)	(.32)	(.31)
14	2.32	2.16	2.29
	(.38)	(.28)	(.3)
15	1.4	1.57	1.95
	(.38)	(.28)	(.3)

Panel B: duration (ms)

speaker ID	object, animate	object, inanimate	subject, animate
1	77.21	69.71	74.2
	(10.61)	(9.88)	(7.10)
2	92.3	91.4	93.15
	(5.76)	(4.36)	(4.67)
3	84.65	74.13	89.96
	(5.76)	(4.35)	(4.67)
4	83.81	70.28	77.3
	(6.55)	(5.25)	(5.58)
5	82.82	77.58	92.87
	(5.77)	(4.39)	(4.67)
6	79.3	72.09	94.88
	(6.38)	(4.67)	(5.58)
7	81.03	81.74	93.7
	(6.76)	(4.94)	(4.86)
8	96.24	93.73	103.86
	(5.83)	(4.33)	(4.67)
9	74.6	69.93	75.74
	(4.45)	(3.22)	(3.77)
10	92.75	91.82	105.78
	(5.83)	(4.48)	(4.78)
11	84.08	75.92	81.25
	(6.12)	(4.42)	(4.78)
12	62.41	64.28	70.97
	(5.77)	(4.45)	(4.67)
13	56.13	56.57	58.32
	(5.90)	(4.90)	(4.78)
14	72.1	68.65	77.41
	(5.90)	(4.35)	(4.67)
15	76.96	73.64	77.38
	(5.83)	(4.39)	(4.71)

Panel C: mean intensity (dB)

speaker ID	object, animate	object, inanimate	subject, animate
1	78.95	78.11	77.84
	(1.31)	(1.22)	(.88)
2	75	76.69	78.15
	(.71)	(.54)	(.58)
3	63.12	62.82	64.4
	(.71)	(.53)	(.57)
4	74.79	75.14	76.58
	(.81)	(.65)	(.69)
5	73.36	75.31	79.14
	(.71)	(.54)	(.58)
6	70.75	71.48	74.3
	(.79)	(.58)	(.69)
7	74.95	74.77	75.73
	(.83)	(.61)	(.61)
8	75.76	76.76	78
	(.72)	(.53)	(.58)
9	76.42	77.85	79.14
	(.55)	(.4)	(.46)
10	79.58	79.96	82.12
	(.55)	(.59)	(.24)
11	79.75	79.18	81.69
	(.76)	(.54)	(.59)
12	77.23	78.55	80.93
	(.71)	(.55)	(.58)
13	79.31	79.52	81.17
	(.73)	(.6)	(.59)
14	75.22	77.6	80.22
	(.73)	(.54)	(.58)
15	77.8	79.22	80.71
	(.72)	(.54)	(.58)

APPENDIX B: *Predicted marginal means for parameters f0 range (panel A), duration (panel B) and mean intensity (panel C) across levels of Word Order*

Panel A: f0 range (ST)

speaker ID	fronted	in-situ	post-posed
1	2.82	2.89	3.24
	(.38)	(.19)	(.38)
2	3.5	2.6	2.14
	(.34)	(.11)	(.33)
3	2.37	2.28	2.61
	(.33)	(.11)	(.32)
4	1.75	1.83	1.83
	(.36)	(.13)	(.35)
5	3.65	2.17	2.25
	(.33)	(.11)	(.33)
6	5.02	3.42	3.33
	(.52)	(.13)	(.49)
7	3.3	2.67	2.62
	(.37)	(.12)	(.34)
8	2.92	2.8	2.8
	(.32)	(.11)	(.32)
9	2.35	2.36	1.71
	(.31)	(.09)	(.3)
10	3.27	2.6	2.95
	(.34)	(.12)	(.32)
11	2.45	1.9	2.36
	(.34)	(.12)	(.31)
12	2.24	1.8	2.18
	(.37)	(.11)	(.35)
13	2.34	2.64	2.24
	(.36)	(.12)	(.38)
14	2.15	2.03	1.95
	(.33)	(.11)	(.31)
15	1.54	1.45	1.9
	(.33)	(.11)	(.32)

Panel B: duration (ms)

speaker ID	fronted	in-situ	post-posed
1	58.12	65.74	77.37
	(5.84)	(2.93)	(5.71)
2	84.39	85.1	93.3
	(5.26)	(1.74)	(5.12)
3	79.28	74.6	86.08
	(5.31)	(1.74)	(5.07)
4	61.74	69.19	82.44
	(5.71)	(2.11)	(5.53)
5	92.53	78.96	96.74
	(5.16)	(1.75)	(5.21)
6	86.16	74.23	74.05
	(8,16)	(2.03)	(7.66)
7	92.31	80.03	91.04
	(5.77)	(1.82)	(5.26)
8	98.07	96.15	95.88
	(5.03)	(1.77)	(5.03)
9	66.92	68.4	77.42
	(4.9)	(1.39)	(4.64)
10	95.56	91.53	93.85
	(5.26)	(1.8)	(5.03
11	72.33	73.59	74.36
	(5.26)	(1.82)	(4.99)
12	64.08	63.06	72.95
	(5.77)	(1.76)	(5.53)
13	51.18	57.67	53.61
	(5.59)	(1.82)	(5.47)
14	62.79	68.35	74.55
	(5.21)	(1.75)	(4.90)
15	70.42	68.31	84.2
	(5.16)	(1.76)	(4.94)

Panel C: mean intensity (dB)

speaker ID	fronted	in-situ	post-posed
1	78.22	75.58	75.34
	(.72)	(.36)	(.7)
2	78.55	76.91	75.69
	(.65)	(.22)	(.63)
3	64.68	63.56	61.93
	(.66)	(.21)	(.62)
4	76.71	75.15	75.01
	(.7)	(.26)	(.68)
5	78.36	75.74	76.21
	(.64)	(.22)	(.64)
6	75.4	71.4	72.2
	(1)	(.25)	(.95)
7	75.86	75.14	74.2
	(.71)	(.23)	(.65)
8	77.08	76.89	73.29
	(.62)	(.22)	(.62)
9	77.08	77.8	76.4
	(.61)	(.17)	(.57)
10	81.97	80.58	78.98
	(.65)	(.22)	(.62)
11	81.9	80.31	79.96
	(.65)	(.22)	(.61)
12	80.23	78.94	76.44
	(.71)	(.22)	(.68)
13	80.66	80.02	79.35
	(.68)	(.69)	(.23)
14	78.4	77.46	77.29
	(.64)	(.22)	(.61)
15	80.69	79.12	79.7
	(.64)	(.22)	(.61)

APPENDIX C: *Predicted marginal means for parameters f0 range (panel A), duration (panel B) and mean intensity (panel C) across levels of IS REF*

Panel A: f0 range (ST)

speaker ID	r-bridging	r-given	r-new
1	1.58	2.99	3.47
	(.35)	(.27)	(.22)
2	2.4	2.69	2.69
	(.24)	(.18)	(.14)
3	2.05	2.89	2.08
	(.23)	(.18)	(.14)
4	1.54	2.21	1.71
	(.28)	(.22)	(.16)
5	2.31	2.43	2.26
	(.24)	(.18)	(.14)
6	3.08	3.76	3.48
	(.31)	(.22)	(.17)
7	2.33	2.82	2.8
	(.25)	(.2)	(.15)
8	2.19	3.18	2.82
	(.24)	(.18)	(.14)
9	2.13	2.73	2.14
	(.2)	(.15)	(.11)
10	2.63	2.79	2.67
	(.24)	(.19)	(.14)
11	1.93	2.17	1.94
	(.24)	(.19)	(.14)
12	1.83	2.04	1.78
	(.24)	(.19)	(.14)
13	2.22	2.9	2.52
	(.26)	(.20)	(.15)
14	1.87	2.41	1.88
	(.24)	(.18)	(.14)
15	1.37	1.68	1.44
	(.24)	(.18)	(.14)

Panel B: duration (ms)

speaker ID	r-bridging	r-given	r-new
1	49.8	65.6	74.4
	(5.22)	(4.15)	(3.44)
2	72.8	93.95	85.64
	(3.68)	(2.85)	(2.18)
3	66.34	84.54	74.67
	(3.66)	(2.85)	(2.18)
4	60.15	78.16	68.75
	(4.31)	(3.43)	(2.55)
5	72.89	87.9	81.51
	(3.68)	(2.86)	(2.18)
6	68.87	81.2	72.86
	(4.79)	(3.38)	(2.61)
7	68.15	85.6	84.95
	(3.86)	(3.04)	(2.25)
8	85.88	93.8	95.7
	(3.68)	(2.88)	(2.2)
9	61.82	77.63	66.5
	(3.07)	(2.35)	(1.75)
10	81.66	99.97	91.25
	(3.8)	(2.94)	(2.22)
11	68.08	79.69	71.93
	(3.8)	(2.95)	(2.23)
12	60.8	68.16	62.62
	(3.76)	(2.89)	(2.23)
13	52.3	58.46	56.59
	(3.88)	(3.03)	(2.26)
14	64.82	75.54	65.67
	(3.68)	(2.86)	(2.18)
15	60.84	75.28	70.41
	(3.68)	(2.86)	(2.19)

Panel C: mean intensity (dB)

speaker ID	r-bridging	r-given	r-new
1	77.78	77.5	76.95
	(.65)	(.52)	(.43)
2	78.2	76.43	76.8
	(.46)	(.36)	(.27)
3	64.43	63.31	66.29
	(.46)	(.35)	(.26)
4	75.94	75.08	75.2
	(.54)	(.43)	(.32)
5	76.97	76.53	78.4
	(.46)	(.36)	(.27)
6	73.12	71.62	71.34
	(.6)	(.42)	(.33)
7	75.07	74.39	82.51
	(.49)	(.38)	(.28)
8	77.51	76.26	76.37
	(.46)	(.36)	(.27)
9	78.4	77.3	77.51
	(.38)	(.29)	(.22)
10	81.34	80.33	80.39
	(.47)	(.37)	(.28)
11	74.26	78.44	80.12
	(.47)	(.37)	(.27)
12	79.96	78.77	78.47
	(.47)	(.36)	(.28)
13	76.41	80.52	84.94
	(.48)	(.38)	(.28)
14	77.86	78.31	76.96
	(.45)	(.36)	(.27)
15	79.7	79.65	79.01
	(.46)	(.36)	(.27)