Preference for single events guides perception in Russian: A phoneme restoration study.
Jesse A. Harris (UCLA) and Natasha Korotkova (Konstanz and Tübingen)

How do language users associate event representations with syntactic strings, when those strings often underdetermine event construal? One possibility is that the processor constructs a single event unless the context or grammar suggests otherwise. While there are many possible origins for a single event preference (e.g., cognitive or representational simplicity) evidence for such a preference has been observed in areas as diverse as collective/distributive ambiguities (Clifton & Frazier, 2012), coordination (Frazier & Clifton, 2013, also Hoeks et al., 2002), quantifier domain ambiguities (Harris et al., 2013), and reciprocals (Fiengo & Lasnik, 1973; Majewski, 2014). We present evidence that the perception of complex predicates in Russian is similarly influenced by a general preference for a single event construal. In essence, perceivers are biased to resolve distorted speech towards a single event interpretation when grammatically licensed.

Serial coverb constructions (see also auxiliary verb constructions) are generally defined as a monoclausal verb complex that shares its arguments and inflectional features like tense and aspect, and are interpreted as single events (Anderson, 2006; Aikhenvald, 2011; Bisang, 2009). Serialization appears in a whole host of languages worldwide, and often alternates with a conjunction, e.g., English *go (and) listen to the radio*. Typically, a restricted set of motion verbs acquires a semantically bleached usage through grammaticalization, though they usually retain both this light and a fully lexicalized use (Traugott & Dasher, 2001; Anderson, 2006; Butt, 2010). Russian is no exception, as a bleached auxiliary use of light verbs of motion (e.g., *idu* 'I walk/go') may immediately precede another verb (*idu slushaju* 'I walk/go listen'), giving rise to unambiguous single event interpretations. As with English and other Germanic languages, an intervening conjunction (e.g., *idu i slushaju* 'I walk/go and listen') is ambiguous between a single and a multiple event interpretation. In the former, the 'going' event is a subpart of the listening event, in which one may initiate an action that results in listening. In the latter, going and listening describe distinct events.

We predicted that when listeners were tasked with resolving an ambiguous or uncertain acoustic signal, they would choose the option which allows a single event structure whenever permitted by the grammar. To test this prediction, we used the phonemic restoration method, in which an intruding cough, tone, or other noise masks an excised phoneme in the acoustic signal, allowing subjects to perceptually "restore" the missing phoneme (Warren, 1970), especially in highly biasing or constraining contexts (Warren & Sherman, 1974; Samuels, 1981). Recent research has shown that perceivers use high-level contextual information in restoration decisions. For example, Stoyneshka et al. (2010) showed that perceivers used prosody in determining how to resolve a masked segment in cases of temporary ambiguity (also Carbary et al., 2015). Another case is Mack et al. (2012), who find that subjects restored a zero subject, e.g., *It* in *(It) seems like it's going well*, when context supported the pragmatic usage preferences for doing so.

In this experiment, we measured the rate at which native Russian speakers restore a conjunction (a single phoneme 'и' *i*) between two verbs, manipulating verb type (Light, Lexical) and Adverb status (Present, Absent) in a crossed 2x2 design, as in (1) below. A single event preference predicts that subjects will restore a conjunction after a Light verb (*idjom;* 'go') less often than after a Lexical verb of motion (*edem*; 'drive'), which necessitates two events, unless an Adverbial (*ne toropjasj*; 'slowly') indicates a separate clause, precluding a serial verb construction.

Twenty-seven native speakers of Russian listened to Russian sentences that were obscured with pink noise in selected locations over headphones in a sound attenuated anechoic chamber. Subjects repeated into a microphone what they thought the message was before it had been distorted, and their responses were recorded. The critical sentences masked * the area between verb phrases, as in (1). As an Internet pilot study (N = 16) revealed that phonetic cues for the conjunction were acoustically present as coarticulation on preceding vowels (also Choi & Keating, 1991), no vowels were included before the conjunction site. Further, materials were equally balanced between cases in which the conjunction was and was not produced in the original recording. All materials were created and produced by a native Russian speaker. In addition to filler sentences, twelve control sentences with narrowly constrained grammatical options for restoration (e.g., case marking) were included to assess speaker competence and attention to the task. Five subjects were removed for failing to accurately restore control items.

The rate of conjunction restoration between verbs was analyzed as a linear mixed effects logistic regression model with sum-coded predictors. Subjects were less likely to restore a conjunction after an initial Light verb (1a) than an initial Lexical verb (1b), t = -2.37. However, this difference only affected cases in which serial verb constructions, and by hypothesis single events, were permitted: subjects restored a conjunction for Lexical verbs more often than for single Light verbs (d = 24%), but not when an Adverbial intervened (d = 2%), t = -2.10. Models that included whether the item was originally produced with a conjunction did not change the results (Fig. 1). An additional post hoc analysis on non-restored items revealed that subjects produced non-significantly longer pause durations after Adverbials following a Light verb, t = 1.6, raising the possibility that subjects also signaled a clause boundary with a prosodic break.

In sum, the rate of conjunction restoration between verbs, and thus perception of an ambiguous or noisy signal, was guided by a general preference for single event interpretation when grammatical possible (Harris et al., 2013; Majewski, 2014). We propose that the bias for single events indicates a preference for representational economy of complex situations in the discourse model, which interacts with grammatical and perceptual decisions. Our results are thus compatible with growing evidence that conceptual constraints restrict how the processor structures bottom-up information, especially for ambiguous input (e.g., Stoyneshka et al., 2010).

(1) **Sample materials.**
a. *Light verb (Adverbial).*
Идем (не торопясь)   * слушаем радио.
*idjom (ne toropjasj)   * slushaem radio*
*walk (neg haste)   * listen   radio*
'We walk/go (slowly) * listen to the radio.'

 b. *Lexical verb (Adverbial).*
 Едем (не торопясь) * слушаем радио.
*edem (ne toropjasj)   * slushaem radio*
*drive (neg haste)   * listen   radio*
'We drive (slowly)   * listen to the radio.'

Fig. 1. Conjunction restoration rate; item originally produced with and without conjunction.