Modelling the interaction of regularity and morphological structure: the case of Russian verb inflection

Selena Rorberi — University of Pisa
Claudia Marzi — Institute for Computational Linguistics - CNR

1 Introduction

Modelling complex inflection systems, such as conjugation in Modern Greek, Italian or Russian, requires careful consideration of a number of factors, ranging from pervasive stem allomorphy to the identification of the appropriate inflection class and the inferential predictability of morpho-phonological processes. Descriptive approaches have taken different views on how to account for degrees of morphological (ir)regularity, while making different predictions about the way speakers process regular and irregular forms in highly-inflecting languages. In the present paper, we assess the psycholinguistic implications of two radically different approaches to the description of the Russian verb system: a more traditional approach dating back to Jakobson (1948), and a Words and Paradigm approach (Brown 1998). Based on recent fMRI evidence (Slioussar et al. 2014) and original results of a neural network simulation with recurrent self-organising maps (Ferro et al. 2011; Marzi et al. 2014; Pirrelli et al. 2015; Marzi et al. 2016), we suggest that both approaches are *prima facie* compatible with Russian data, while being in contrast with Pinker's claim that the regular-irregular distinction is an epiphenomenon of the storage-processing dichotomy in the human language faculty (Pinker & Ullman 2002). We argue that this evidence lends support to integrative models of the mental lexicon (Marzi & Pirrelli 2015), accounting for a graded interaction between regularity and morphological structure.

2 The Russian verb

Traditionally, the description of Russian verb inflection is cast into the classical derivational analysis first proposed by Roman Jakobson (1948). For each verb lemma, Jakobson postulates the existence of a unique, underlying stem, which may undergo a variety of morpho-phonological processes as a function of the class to which the verb belongs, and its specific set of endings (including class-specific thematic vowels, or "thematic ligatures"). A number of verb classes are identified, whose variety reflects the type and number of the morpho-phonological processes needed to turn an underlying base into a surface allomorph. Classes are identified by the suffix classifier in the verb stem*: -aj-*, *-ej-*, *-a-*, *-e-*, *-i-*, *-o-*, *-ova-*, *-avaj-*, *-nu-*. In particular, the classifier determines the conjugation class (i.e. the specific set of inflectional endings selected by the verb), the adjustment of the root final consonant (i.e. the root consonant immediately preceding the classifier), and the suffix alternation (e.g. *-ova-* alternates with *-uj-*). To illustrate, the stem *chitaj-* of the verb *chitat'* ('read') drops the final *-j* before an ending that begins with a consonant (e.g. past tense *chita-l*), but keeps *-j* when the ensuing ending begins with a vowel (e.g. *chitaj-u* 'I read'). In contrast, the stem *pisa-* of the verb *pisat'* ('write') drops its final vowel when the affixed ending begins with a vowel. In turn, this triggers consonant softening throughout the present indicative paradigm (e.g. *pish-u* 'I write'). As verb stems in any class are assumed to undergo some stem alternation, regularity is measured by the number of applicable processes. Accordingly, the *-aj-* class is more regular than the *-a-* class, since the former undergoes consonant truncation only, whereas the latter undergoes both vowel truncation and consonant softening.

More recently, Brown (1998) proposes a paradigm-based account of Russian verb inflection, cast into a Network Morphology framework (Brown & Hippisley 2012). The analysis focuses on the number of stem alternants associated with specific cells in a verb paradigm, independently of the degree of formal predictability or the number of processes involved in stem formation. Unlike Jakobson's analysis, in a verb like *chitat'* the stem is analysed as ending in a vowel (*chita-*); *-j-* is infixed when the stem is followed by a vowel-initial ending (*chita-j-u*). The so-called *-aj-* class in fact includes those verbs that keep their *a*-ending stem unaltered throughout the whole paradigm. This represents a kind of default class. In contrast, a verb paradigm with more stem alternants is less regular and more difficult to master and generalise than a verb paradigm with fewer or no stem alternants. Thus, regularity is expressed in terms of surface relations between paradigmatically-related verb forms. In regular paradigms, invariant stems are shared by all inflected forms, and are transparently perceived by the speakers. Conversely, irregular paradigms select more than one stem alternant, which are differently indexed, depending on the verb class.

2.1 Psycholinguistic implications

In spite of considerable differences in their formal apparatus, both approaches account for a graded notion of morphological regularity and its interaction with word processing. Following Jakobson, the more processes are involved in mapping allomorphs onto an invariant stem, the longer it takes a speaker to master them. In Brown's account, paradigms with more stem alternants are more difficult to process because their simultaneous availability in the speakers' long-term memory causes their co-activation and mutual competition during processing. For example, competing co-activation of *pisa-* and *pish-* as stem alternants of *pisat'* slows down their processing in recognition.

In general, paradigms with more stem alternants require stipulation of more morpho-phonological processes. Due to this correlation, the most regular class of *-aj-* verbs in Jakobson's approach (requiring one *j*-deletion rule) coincides with the regular class of invariant verb stems in Brown's account. In addition, both accounts predict that difficulty of processing, as well as ease of generalisation and learning, should vary continuously as a function of graded levels of regularity. Nonetheless, there is one point where the two accounts diverge. In radically amorphous versions of the *Word-and-Paradigm* approach (Blevins 2016), as well as in connectionist frameworks, the mapping of an input inflected form onto its sublexical constituents ((prefix +) stem + ending, for Russian verb forms) is a continuous function of the statistical regularities of inflectional paradigms. Accordingly, perception of morphological boundaries may vary as a result of the probabilistic support sublexical boundaries receive from frequency distributions of surface exemplars (e.g. Hay & Baayen 2005; Plaut & Gonnerman 2000; Rueckl & Raveh 1999). It follows that processing of regularly inflected forms should be more sensitive to their morphological structure and to type frequency effects than the processing of irregulars. Conversely, irregulars are processed holistically, in a way that is sensitive to token frequency effects.

3 Computational evidence

We provide data-driven evidence of the complex interaction in processing of a graded notion of (ir-)regularity and the morphological structure. Sixteen fully inflected verb forms have been selected for each of the 50 top frequency Russian verb paradigms (i.e. 50 aspectual pairs, which include 10 present and past tense imperfective forms, 6 perfective forms for the future tense) sampled from a reference corpus (Jakubíček et al. 2013). Without any information of morphological structure, they are learned by a recurrent self-organising neural network

(TSOM), consisting of a two-dimensional grid of artificial memory/processing nodes that dynamically memorise input strings as chains of maximally-responding processing nodes (Best Matching Units). The prediction-driven bias of its temporal layer of re-entrant connections makes strong expectations over upcoming symbols accounting for successful serial word processing. Figure 1 illustrates the dynamic of word access at the end of learning (i.e. epoch 100) by showing prediction rates at each letter position relative to the stem-ending boundary (or morpheme boundary, centred on x-axis = 0). Prediction scores are calculated by incrementally assigning each correctly anticipated symbol in the input a 1-point score. The more input symbols are anticipated, the easier the prediction of the verb form, the lower its processing load.

Our evidence suggests that perception of morphological structure interacts with regularity and formal transparency. The more prominent increase in prediction rates on more regular stems suggests a clear paradigmatic effect: the more verb forms share the same stem, the easier their processing. On the contrary, the drop in prediction reflects an increase in the processing effort made by the map in predicting an upcoming inflectional ending at the end of the stem. We take such a discontinuity to mark a clear structure-driven effect of processing "surprisal" (Levy 2008), due to an increase in entropy of the transitional probability from a regular stem to its grammatical endings. This is confirmed by the steeper increase in prediction rates for inflectional endings (positive x values in Figure 1) when they follow more irregular stems: stem allomorphs can anticipate inflection information thus reducing uncertainty for selection of ending.

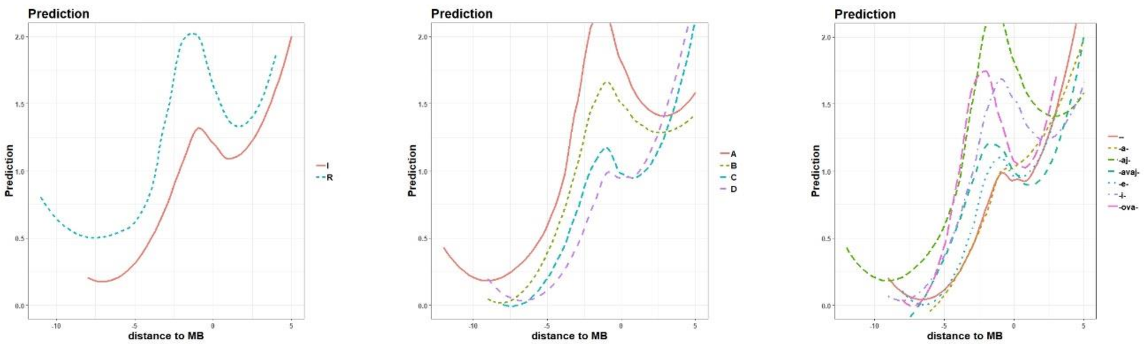

Figure 1. Regression plots of interaction effects between morphological (ir)regularity and distance to morpheme boundary (MB), in non-linear models (GAMs) fitting the number of symbols predicted by a TSOM: categorical fixed effect are (left panel) regularity (green dashed lines) vs. irregularity (red solid lines), (central panel) a gradient of regularity, and (right panel) suffixes classes.

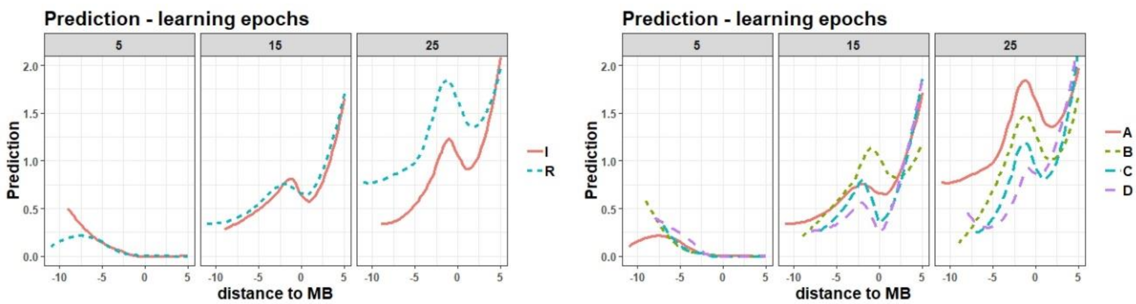

Figure 2. Regression plots of interaction effects between morphological (ir)regularity and distance to morpheme boundary (MB) for learning epochs 5, 15, 25, in non-linear models (GAMs) fitting the number of symbols predicted by a TSOM: categorical fixed effect are (left panel) regularity (green dashed lines) vs. irregularity (red solid lines), and (right panel) a gradient of regularity.

The effect is consistent. We observed it using three classification criteria for inflection regularity of different grain-size: (a) the traditional dichotomy between the class of *-aj-* verbs (Regular) and the class of non *-aj-* verbs (Irregular); (b) a more granular subdivision between *-aj-* verbs (class A), productive *-i-* and *-ova-* verbs (class B), *-a-*, *-e-*, *-avaj-* verbs (class C), and radically suppletive paradigms (class D); and (c) all suffix-based classes attested in our training set. In all cases, more regular verb classes, when compared with less regular classes, show higher prediction rates overall, while exhibiting a greater discontinuity in prediction at the stem-ending boundary. As shown in Figure 2, perception of morphological structure gradually emerges through the training epochs, as learning progresses.

Our results well agree with evidence of word processing load reported by Slioussar and colleagues in the task of generating 1Sg present tense forms of regular and irregular Russian verbs (Slioussar et al. 2014). In their experiment, regulars were found to require less attention, working memory and decision-making than irregulars. While their evidence appears to support an integrative model of word processing, our results address the important, related question of how similar effects may arise in a recurrent self-organising network that simulates the concurrent dynamic storage of paradigmatically related forms. In the end, the interaction between regularity and morphological structure appears to be more compatible with a *Word-and-Paradigm* account of Russian verb inflection, than with Jakobson's account, which does not make the same prediction.

References

Brown, Dunstan. 1998. Stem Indexing and Morphological Selection in the Russian Verb: a Network Morphology Account. In Ray Fabri, Albert Ortmann & Teresa Parodi (eds.), *Models of Inflection*, 196-224. Max Niemeyer Verlag.

Brown, Dunstan & Andrew Hippisley. 2012. *Network Morphology: A Defaults-based Theory of Word Structure*. Cambridge University Press.

Ferro, Marcello, Claudia Marzi & Vito Pirrelli. 2011. A Self-Organizing Model of Word Storage and Processing: Implications for Morphology Learning. *Lingue e Linguaggio*, X(2). 209-226.

Jakobson, Roman. 1948. Russian Conjugation. *Word*, 4(3). 155-167.

Jakubíček, Miloš, Adam Kilgarriff, Vojtêch Kovář, Pavel Rychlý & Vít Suchomel. 2013. The tenten corpus family. In *7th International Corpus Linguistics Conference CL*, 125-127. https://www.sketchengine.eu/rutenten-russian-corpus/

Levy, Roger. 2008. Expectation-based Syntactic Complexity. *Cognition* 106. 1126-1177.

Marzi, Claudia, Marcello Ferro & Vito Pirrelli. 2014. Morphological structure through lexical parsability. *Lingue e Linguaggio*, XIII(2). 263-290.

Marzi, Claudia, Marcello Ferro, Franco Alberto Cardillo & Vito Pirrelli. 2016. Effects of frequency and regularity in an integrative model of word storage and processing. *Italian Journal of Linguistics*, 28(1). 79-114.

Pinker, Steven & Michael Ullman. 2002. The past and future of the past tense. *TRENDS in Cognitive Sciences* 6(11). 456-463.

Pirrelli, Vito, Marcello Ferro & Claudia Marzi. 2015. Computational complexity of abstractive morphology. In Baerman, Matthew, Dustan Brown & Greville Corbett (eds.), *Understanding and Measuring Morphological Complexity*. Oxford: Oxford University Press. 141-166.

Slioussar, Natalia, Maxim V. Kireev, Tatiana V. Chernigovskaya, Galina V. Kataeva, Alexander D. Korotkov & Svyatoslav V. Medvedev. 2014 An ER-fMRI study of Russian inflectional morphology. *Brain & Language* 130. 33-41.