

Studies to examine native dispreference for grammatical subextraction in Russian

Jonathan L. Rossiⁱ

November 11, 2013

1. Introduction

In ‘*Native versus heritage processing of subextraction in Russian*,’ we examined a trend in recent work, which indicates that heritage speakers lack the language abilities of native speakers in domains where narrow syntax interfaces with grammar-external cues. We presented supporting data from a study of heritage versus native acceptance of Russian subextraction. Our results showed that both speaker groups were able to classify ungrammatical subextraction as being unacceptable, while only the heritage group rated the grammatical subextraction acceptable. The native group, meanwhile, exhibited a dispreference for the grammatical subextraction. We argued that this dispreference is due to the grammar-external information inherent in the subextraction, and that heritage speakers do not share the dispreference because they lack the native speakers’ ability to utilize this information in making acceptability judgments.

The current paper proposes methods to further investigate exactly what information the native speakers are accessing that causes them to rate the grammatical sentences unacceptable. Carrying out such an investigation will, first, help us better understand how our native group is processing subextraction. Second, it will further our knowledge of heritage speaker grammars (we argued in the previous paper that whatever native speakers are using to make judgments on the grammatical subextraction, heritage speakers most likely lack). Finally, it will set us up to reexamine, under a narrower lens, the models of language acquisition we proposed in the previous paper that aim to explain the weaker heritage abilities with respect to grammar-external domains.

2. Hypotheses

We argued previously that the intonation patterns involved in subextraction index extra-grammatical information, and suggested two ways this information might cause the native speaker dispreference for grammatical subextraction. One possibility is that the subextraction stimuli we used do not offer an acceptable intonation pattern and thus native speakers, expecting the subextraction to occur with intonation patterns that correlate with contrastive topic and contrastive focus, judged the sentences to be unacceptable (*hypothesis 1*). Another possibility is that, because we did not present contextual information that might license the subextraction construction, native speakers judged the sentences to

ⁱ Supervised by Maria Polinsky and Annie Gagliardi

be infelicitous (and therefore unacceptable) regardless of their syntax or prosody (*hypothesis 2*). In the following sections, we will outline methods for how to determine the validity of each of these hypotheses, which we should note are not mutually exclusive. By carrying out the studies proposed, we will better understand the role of intonation patterns versus contextual information in the native dispreference for grammatical subextraction.

2.1 Hypothesis 1

Pereltsvaig² claims that native speakers expect sentences involving subextraction to express contrastive topic or contrastive focus, each of which is typically accompanied by a specific prosody pattern. Contrastive topic usually corresponds to the *Intonational Construction 5* (IC-5) and, contrastive focus, the *Intonational Construction 2* (IC-2), as discussed in the next section. *Hypothesis 1* posits that the native speaker group disprefers the grammatical subextraction items in our study because they do not consistently use the appropriate IC-5 and IC-2 intonation patterns. *Study 1* is designed to test whether this is the case.

Study 1

IC-5 is characterized by a rising peak on the contrastive topic element and a falling peak on the subsequent contrastive focus element, with a high plateau in between. Both peaks are accompanied by strengthened lexical stress. IC-2, on the other hand, simply has a falling peak with strengthened lexical stress on the contrastive focus element.³ The goal of this study is to determine whether our stimuli items have pitch contours that resemble IC-5 and IC-2.

Using acoustic analysis software, such as Praatⁱⁱ, we will generate pitch contours of the fifty-six grammatical stimuli items. For our purposes, it will be sufficient to visually check these contours for the rising and falling peaks of IC-5 and IC-2. If we find a correlation between the pitch contours of the stimuli and the contours of IC-5 and IC-2, we can conclude that the reason the native group exhibits a dispreference for these items is not likely due to a lack of appropriate intonation patterns. We should, however, be wary of ruling out this hypothesis altogether, as it is possible both *hypothesis 1* and *hypothesis 2* could be at play in some capacity.

If we find no such correlation, however, we must investigate whether the lack of appropriate intonation could be the reason for the dispreference. Using fresh recordings, we would need to create a new group of stimuli items that use the old intonation patterns, as well as another group of items that conform to the IC-5 and IC-2 intonation patterns. We would then test the old intonation patterns on one new group of participants, and the IC-5 and IC-2 intonation patterns on another, separate group of participants. If we found significantly higher native speaker acceptance ratings of the IC-5 and IC-2 sentences (using an ANOVA to look for a significant effect of *intonation pattern*), we could safely say that the intonation patterns indeed contributed to the original native speaker dispreference for grammatical subextraction. If we found no significant difference between native acceptance ratings for the old intonation patterns and the new, we would need to explore other possible explanations for the dispreference.

ⁱⁱ <http://www.fon.hum.uva.nl/praat/>

2.2 Hypothesis 2

A second possible reason for low native acceptability ratings is a lack of contextual information to license the use of subextraction. *Hypothesis 2* suggests that, regardless of whether our stimuli employ the appropriate syntax and prosody, our native speakers found them infelicitous due to the fact that we did not provide context that might elicit a sentence involving contrastive topic or contrastive focus. *Study 2* offers a method for investigating this claim.

Study 2

We will rerun the study with the grammatical stimuli, using the same items, adding contextual information to precede each of the sentences. This contextual information will fall under two conditions: 1) context that licenses subextraction (*informative context condition*) and 2) context that does not license subextraction (*uninformative context condition*). Half of the stimuli will be accompanied by information that meets the *informative condition*, and the other half, information that meets the *uninformative condition*. The following example shows a possible *informative* and *uninformative* context for the subextraction item ‘*It is possible to buy STRAWBERRY jam!*’

- | | |
|--------------------------|--|
| (<i>informative</i>) | <i>Amy is unaware that the corner store sells only strawberry jam. She asks the clerk whether it is possible to buy blueberry jam.</i> |
| (<i>uninformative</i>) | <i>The corner store sells many types of jam. Amy asks the clerk what kind of jam she can buy.</i> |

We will then use an ANOVA on the native speaker acceptance ratings for the *informative* items and the *uninformative* items to see whether there is a significant effect of *condition*. If we find that ratings for the *informative condition* are significantly higher than ratings for the *uninformative condition*, we can safely say that *hypothesis 2* is correct (i.e., that the lack of context in our study did have a negative effect on our native group’s ratings). If, however, there is no statistically significant difference in ratings between the two conditions, we cannot say for certain that the lack of context was not the reason for the native dispreference. While this is certainly a likely possibility, it may also be the case that our conditions were not designed carefully enough, or that there are other factors affecting the results of which we are not aware.

3. Discussion

It is clear from our data, and the discussion in the previous paper, that our native speakers do exhibit a dispreference for the grammatical subextraction stimuli that is based on grammar-external information. We argued previously that the dispreference likely comes from expectations about intonation patterns associated with subextraction. However, we were unable to say where exactly this expectation came into play. Native speakers either expected the subextraction stimuli to employ IC-5 or IC-2, and did not have this expectation fulfilled, or they perhaps lacked an expectation that they would hear subextraction without prior context to license such a construction.

By carrying out the above studies, we can better pinpoint where the expectation is causing the dispreference. Clarifying exactly where this dispreference happens will be interesting in its own right, since it will give us insight into exactly how native speakers are processing the subextraction construction.

We will also infer something about the abilities of the heritage speakers in our study. We argued in the previous paper that whatever information is causing the native dispreference must be the very same information that heritage speakers are unable to access. Therefore, if we can say precisely what knowledge our native group is utilizing in rating the grammatical subextraction, we can say precisely what knowledge the heritage group lacks.

Finally, we can use the results of the studies proposed in this paper to comment on Russian heritage speaker grammars in general. In the previous paper we discussed two different models of heritage language acquisition. The *nativism* model suggested that grammar learning is privileged, allowing heritage speakers to learn acceptable Russian syntax before picking up on extra-grammatical constraints. Meanwhile, according to the *data-driven* model, heritage speakers track data about syntax and grammar-external constraints as they are exposed to it during acquisition, though it is perhaps the case that information about syntax is trackable at an earlier stage of development than information about extra-grammatical cues.

What can *study 1* and *study 2* tell us about the validity of each of these acquisition models? If the first study reveals that *hypothesis 1* is true (i.e., that native speakers rated the grammatical subextraction low because our stimuli did not use the expected IC-5 or IC-2 intonation patterns), it follows that our native speakers must know which particular prosody patterns are and are not acceptable with subextraction. This implies, as we have argued previously, that heritage speakers must in turn lack this knowledge. In other words, they have acquired the knowledge of syntax but not of the prosody patterns that can accompany a particular syntactic construction. This aligns quite well with the *nativism* model (syntax is learned because grammar learning is privileged but knowledge of prosody patterns, which are external to syntax, is not acquired as efficiently). The *data-driven* model, however, no longer seems as plausible: if heritage speakers are simply tracking data as they are exposed to it, why should they acquire knowledge of syntax but not of prosody patterns when prosody patterns are so closely linked to syntax (much more so than, say, discourse)?

If *hypothesis 2* is true (i.e., that native speakers rated the grammatical subextraction low because they judged it to be infelicitous without context to license it), it follows that our native speakers must know that subextraction is not acceptable unless accompanied by particular contextual information. Our heritage speakers, then, must lack this knowledge. That is, they can identify the proper syntax for subextraction but are unaware of which contextual situations license it, or even that contextual information is necessary at all. Again, this fits nicely with the *nativism* model (syntax learning is privileged and knowledge about context is acquired over a longer period of exposure).

The *data-driven* model may also look appealing. Knowledge about context is in the realm of discourse, far external to syntax and prosody, and so perhaps it is more difficult to track until an advanced stage of development. Still, some might argue, why should knowledge of this requisite context not be acquired with knowledge of the syntax if the contextual information and syntax always occur together? In order to address this question, we would need to design a study to examine whether it is indeed easier for heritage speakers to track data about syntax than data about grammar-external constraints like contextual relationships at various points in their language development.

4. Conclusion

The studies we have proposed will, first, further our understanding of how native speakers process subextraction. Second, they will allow us to better identify exactly what knowledge the heritage speakers in our study are lacking. Finally, they will set the stage for further discussion about the merits of different models of heritage acquisition, such as the *nativism* and *data-driven* models. Although much more work is required before we can begin to rule out particular theories of heritage language acquisition in search of the correct model, the results of these studies will bring us closer to understanding the necessary steps of knowledge acquisition each theory requires.

References

¹ Rossi, Jonathan L. “Native versus heritage processing of subextraction in Russian.” (Unpublished paper, Polinsky Language Sciences Lab, Harvard University, 2013).

² Pereltsvaig, Asya. “Split Phrases in Colloquial Russian.” *Studia Linguistica* 62:1 (2008): 5–38.

³ Ibid.