Poster Script: Embeddings in monolingual and heritage Russian

PhD: Maria Martynova

PIs: Luka Szucsich & Natalia Gagarina

P10* | Humboldt-Universität zu Berlin & Leibniz-Zentrum Allgemeine Sprachwissenschaft

Main finding. Bilingualism has no impact on syntactic complexity in Russian: heritage speakers produce as much embeddings as monolinguals do.

Intro. Heritage languages are particularly open to linguistic variation (Montrul 2016). In some areas, such as syntax, heritage speakers (HSs) often show noncanonical patterns (Polinsky 2018). We want to look at syntactic complexity as a possible domain of dynamicity in heritage languages. Embedded (subordinate) clauses are often considered to indicate syntactic complexity (Sánchez Abchi & De Mier 2017 for heritage Spanish; Levy et al. 2013, for relative clauses in monolingual Russian, Lintunen & Mäkilä 2014 for L2 English by Finnish monolingual speakers). For instance, a high number of embeddings is associated with a high degree of syntactic complexity (s. ibid.). According to Karlsson (2007:366), "[t]he notion EMBEDDING refers to all types of clauses occurring as subordinate parts of their superordinate clauses (which themselves may be either main or subordinate)." Previous research on embeddings/ syntactic complexity under language contact suggests differences between bilingual and monolingual speakers (Pavlenko 2000, Treffers-Daller et al. 2008, Schmid 2010). One possible mechanism behind this discrepancy might be the concept of simplification or reduction (Dussias & 2004, Pavlenko 2010, Ribbert & Kuiken 2010). Also, embedded clauses are acquired later than main clauses (Kuiken & Vedder 2019, for L2 acquisition, Cejtlin 2000 on acquisition of syntax by monolingual Russian-speaking children, Ovčinnikova 2011, on monolingual Russian and heritage Russian in Israel, Čirševa 2011 on Russian-English simultaneous bilinguals) and later acquired structures are potentially prone for attrition (Slobin 1986). However, there are only few studies investigating embeddings in bilinguals and their results are somewhat contradictory. For instance, a study conducted by Treffers-Daller et al. (2008) reports that Turkish-German bilinguals (heritage speakers and returnees) produce less embeddings in their narrations as compared to monolingual Turkish speakers. In contrast, a study run by Yılmaz (2011) concludes that (sequential) Turkish-Dutch bilinguals and monolingual Turkish speakers produce equal number of subordinate clauses in their L1 Turkish. To our knowledge there has been no research on embedding rate in productions of HSs of Russian. This study aims to fill this gap.

RQ: Which factors contribute to the production of more complex clauses in terms of a higher rate of embeddings in narrations of monolingual and bilingual heritage Russian speakers?

Hypothesis: Based on the literature, HSs of Russian are expected to produce less embeddings as compared to monolinguals or they will tend to avoid them, since embedded clauses are acquired later (Kuiken & Vedder 2019, Cejtlin 2000, Ovčinnikova 2011, Čirševa 2011) and might be therefore prone to reduction (Dussias & 2004, Pavlenko 2010, Ribbert & Kuiken 2010). Hence, we expect the factor bilingualism to have an impact on embeddings.

Method. We tested our RQ on the data of 20 monolingual speakers of Russian, 36 bilingual HSs of Russian in Germany and 55 bilingual HSs of Russian in the US. Each speaker produced 4 documents in 4 different language situations (formal spoken, formal written, informal spoken, informal written). The data were driven from the RUEG-RU_0.4.0 corpus (Wiese at al. 2020) and contained manual annotation of clause type, which was subsequently merged into two levels: main vs. embedded

clauses. Hence, we included the following factors, namely bilingualism (bilingual vs. monolingual speakers), number of main clauses per document, formality (formal vs. informal) and mode (spoken vs. written) in our analysis, since the latter belong to the RUEG communicative situations experiment design (Wiese 2020). To account an inter-individual speaker variation, we applied linear mixed-effects modelling using R package lme4 (Bates et al. 2015). The dependent variable was the number of embedded clauses per document, the independent variables were the number of main clauses per document, bilingualism, formality and mode.

Results. The results showed no main effect of bilingualism and also no interaction of bilingualism with other factors. However, formality and mode seem to have impact on the number of embeddings across the groups. Specifically, formal and spoken communicative situations involve more embeddings as compared to informal and written communicative situations in our data sample. Interestingly, differences in the degree of syntactic complexity across different communicative situations are more pronounced in monolingual than in bilingual speakers of Russian. Also, the number of main clauses positively correlates with the number of embedded clauses.

Figure 1: Mean percent of embeddings by monolingual and bilingual (HSs) speakers of Russian, error bars represent SE.
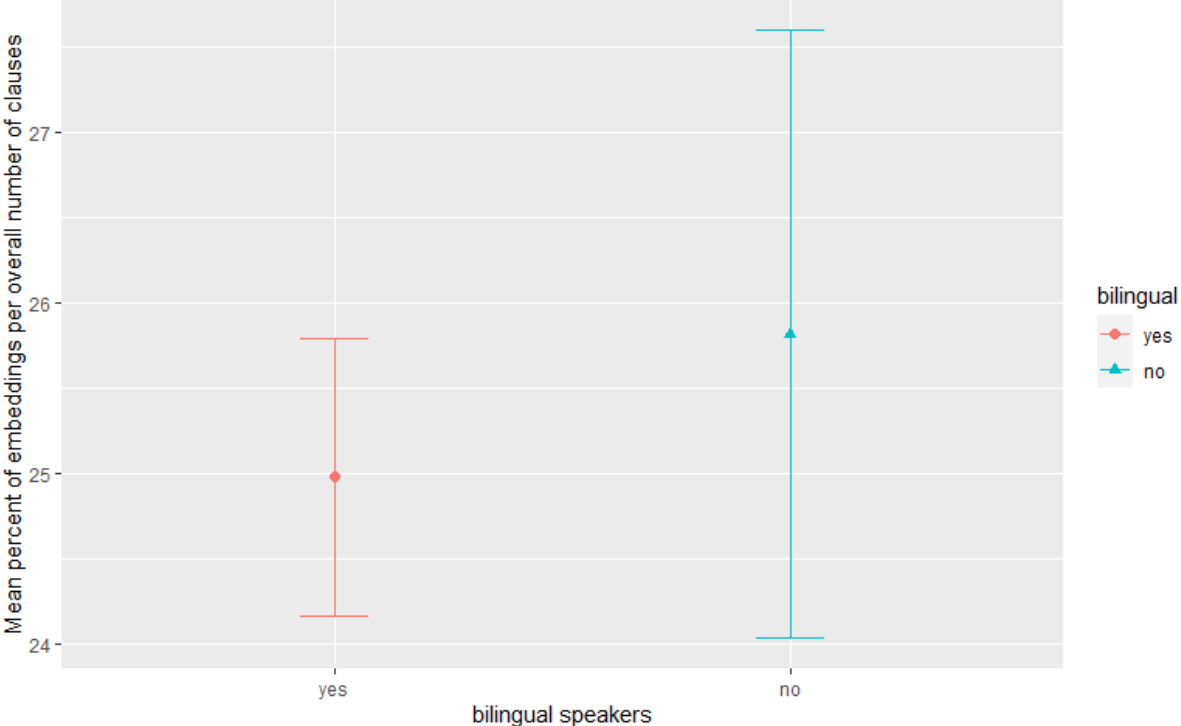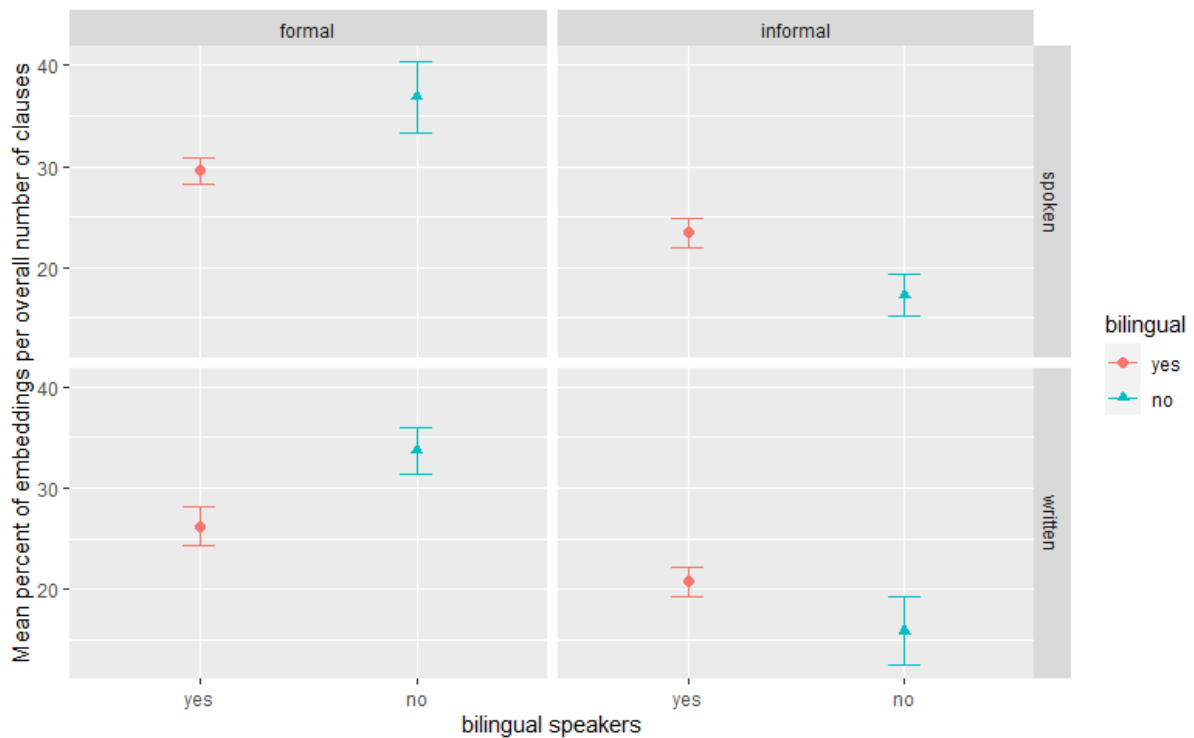
Figure 2: Mean percent of embeddings by monolingual and bilingual (HSs) speakers of Russian grouped by communicative situations, error bars represent SE.

Discussion & Conclusion. The study aimed at investigating whether bilingualism and other factors such as formality, mode and the number of main clauses play a role in syntactic complexity in narrations in Russian. The results showed that participants across the groups behaved similarly - monolinguals did not quantitatively differed from the bilingual HSs of Russian. This finding reveals no impact of bilingualism on syntactic complexity. Instead, the number of main clauses, formality and mode seem to have an influence on the number of embeddings in Russian. That means, that the degree of syntactic complexity in terms of number of embeddings is dependent on the communicative situation. Besides, syntactic complexity is not exclusively a feature of the production of monolingual speakers.

Model description

Model formula:

mem.01f <- lmer (emb ~ main + bilingual + formality + mode + (1|speaker), data=main)

The interactions of fixed effects were not included since the model with interactions were not significantly better than the model without interactions.

```
Fixed effects:
                  Estimate Std. Error        df t value Pr(>|t|)
(Intercept)        1.54649    0.52346 380.39797   2.954  0.00333 **
main               0.34738    0.03045 429.50548  11.407  < 2e-16 ***
bilingualno       -0.06013    0.56883 111.44140  -0.106  0.91600
formalityinformal -1.45363    0.24254 329.01799  -5.993  5.4e-09 ***
modewritten       -0.78094    0.29646 389.76415  -2.634  0.00877 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

References

Bates, D., Maechler, M., Bolker, B. & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software, 67(1), 1-48.

Cejtlin, S. (2000). Jazyk i rebenok: Lingvistika detskoj reči: Učeb. Posobie dlja stud. vysš. učeb, zavedenij [Language and the child: Linguistics of children's speech: A study guide for university students]. VLADOS.

Čirševa, G. (2011). Osobennosti odnovremennogo usvoenija russkoj i anglijskoj grammatiki [Features of simultaneous acquisition of Russian and English grammar]. In Cejtlin, S., & Eliseeva, M. B. (Eds.), Put' v jazyk: Odnojazyčie i dvujazyčie [Monolingual and bilingual path to language] (pp. 202-219). Jazyki slavjanskix kul'tur.

Dussias, P.E. (2004). Parsing a first language like a second: The erosion of L1 parsing strategies in Spanish-English bilinguals. International Journal of Bilingualism, 8(3), 355–371.

Kuiken, F., & Vedder, I. (2019). Syntactic complexity across proficiency and languages: L2 and L1 writing in Dutch, Italian and Spanish. International Journal of Applied Linguistics, 29(2), 192-210.

Levy, R, Fedorenko, E., & Gibson, E. (2013). The syntactic complexity of Russian relative clauses. Journal of Memory and Language, 69, 461–495.

Lintunen, P., & Mäkilä, M. (2014). Measuring Syntactic Complexity in Spoken and Written Learner Language: Comparing the Incomparable? Research in Language, 12(4), 377-399.

Montrul, S. (2016). The acquisition of heritage languages. Cambridge: Cambridge University Press.

Karlsson, F. (2007). Constraints on multiple center-embedding of clauses. Linguistics 43, 365–392.

Ovčinnikova, I. (2011). O stanovlenii kategorii taksisa (na materiale povestvovanij doškol'nikov) [On the formation of the taxis category (based on the narratives of preschoolers)]. In Cejtlin, S., & Eliseeva, M. B. (Eds.), Put' v jazyk: Odnojazyčie i dvujazyčie [Monolingual and bilingual path to language] (pp. 122-134). Jazyki slavjanskix kul'tur.

Pavlenko, A. (2000). L2 influence on late L1 bilingualism. Issues in Applied Linguistics, 11(2), 175–205.

Pavlenko, A. (2010). Verbs of motion in L1 Russian of Russian-English bilinguals. Bilingualism: Language and Cognition, 13(1), 175–205.

Polinsky, M. (2018). Heritage languages and their speakers. Cambridge: Cambridge University Press.

R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing. https://www.R-project.org/

Ribbert, A. & Kuiken, F. (2010). L2-induced changes in the L1 of Germans living in the Netherlands. Bilingualism: Language and Cognition, 13(1), 41–48.

Sánchez Abchi, V., & De Mier, V. (2017). Syntactic complexity in narratives written by Spanish heritage speakers. Vigo International Journal of Applied Linguistics, 14, 125-148.

Schmid, M.S. (2010) Languages at play: the relevance of L1 attrition to the study of bilingualism. Bilingualism: Language and Cognition, 13(1), 1–7.

Slobin, D. (1986). The acquisition and use of relative clauses in Turkic and Indo-European languages. In D.I. Slobin & K. Zimmer (Eds.), Studies in Turkish linguistics, 273–294. Amsterdam: John Benjamins.

Treffers-Daller, J., Özsoy, A.S. & van Hout, R. (2007). (In)Complete acquisition of Turkish among Turkish-German bilinguals in Germany and Turkey: an analysis of complex embeddings in narratives. International Journal of Bilingualism and Bilingual Education, 10(3), 248–276.

Wiese, H. (2020). Language Situations: A method for capturing variation within speakers' repertoires. In Asahi, Y (Ed.), Methods in Dialectology XVI, 105-117. Peter Lang.

Wiese, H., Alexiadou, A., Allen, S., Bunk, O., Gagarina, N., Iefremenko, K., Jahns, E., Klotz, M., Krause, T., Labrenz, A., Lüdeling, A., Martynova, M., Neuhaus, K., Pashkova, T., Rizou, V., Tracy, R., Schroeder, C., Szucsich, L., Tsehaye, W., Zerbian, S. & Zuban, Yulia. (2020). RUEG Corpus (0.3.0) [Data set]. Zenodo. https://doi.org/10.5281/zenodo.3765218

Yilmaz, G. (2011). Complex embeddings in free speech production among late Turkish-Dutch bilinguals. Language, Interaction and Acquisition 2(2), 251–275.