Bernhard Brehmer*, Dominika Steinbach and Vladimir Arifulin

Heritage languages and the 'multilingual boost': intercomprehension skills of Russian and Polish heritage speakers in Germany

Abstract: The article focuses on whether and to which extent heritage bilinguals make use of their heritage language while developing receptive skills in unknown languages which are either related to the majority language or the heritage language. Thirty four adolescent heritage speakers of Russian and Polish and a control group of thirty three German monolinguals were first exposed to a text in Swedish. The monolingual control group was matched with regard to age, educational background, foreign languages learned at school as well as proficiency in English. All participants had to determine the parts of speech of ten items from the text, translate them into German, and extract the main pieces of information from the text. In a second step, the heritage speakers completed a similar task with an unknown Slavic target language (Serbian). The results revealed no bilingual advantage of the heritage speakers over the monolinguals in the Swedish task. Furthermore, they scored lower in the Serbian trial. We treat this as evidence that access to the heritage language as a resource for solving these tasks is limited compared to the majority language and English which might be due to lesser metalinguistic knowledge about structures of the heritage language.

Keywords: bilingual advantage; heritage languages; intercomprehension; interlingual transfer; receptive multilingualism

1 Introduction

An increasing number of studies has been published on bilinguals' linguistic and cognitive abilities compared to those of monolinguals over the last decades.

Corresponding author: **Bernhard Brehmer, Linguistics, University of Konstanz, Konstanz, Germany, E-mail: bernhard.brehmer@uni-konstanz.de*
Dominika Steinbach, Linguistics, University of Konstanz, Konstanz, Germany,
E-mail: dominika.a.steinbach@gmail.com
Vladimir Arifulin, Slavic Department, University of Greifswald, Greifswald, Mecklenburg-Vorpommern, Germany, E-mail: vladimir.arifulin@uni-greifswald.de

Facilitative effects of multilingualism are typically identified in the realm of particular cognitive skills like executive control. However, a systematic review of the relevant literature reveals a number of studies that cannot confirm overall superior performance of bilinguals over monolinguals in different cognitive control tasks (Adesope et al. 2010; Gathercole et al. 2014; Paap et al. 2018). Furthermore, Adesope et al. (2010) mention the possibility of a "publication bias" against studies showing a lack of a bilingual advantage on tasks targeting cognitive performance. The second realm where a bilingual advantage has been regularly claimed is L3 acquisition (see Cenoz 2003 for an overview): Facilitative effects of bilingualism on the acquisition of additional languages are generally linked to a higher metalinguistic awareness of bilinguals over monolinguals and more profound experiences in language acquisition (Bialystok 2001; Jessner 2008). However, the outcomes of studies dealing with a proposed "multilingual boost effect" (Berthele and Udry 2019) in L3 acquisition by heritage bilinguals are again far from being unequivocal. While some studies on the acquisition of English or Romance foreign languages at school yielded supportive evidence for a bilingual advantage, others did not (see Lorenz et al. 2020 for a recent overview). In light of this confusing picture, our contribution provides new evidence on potential bilingual advantages in contexts of additional language learning. For our own research, we will limit the scope of inquiry to the acquisition of receptive multilingual skills.

2 Multilingual repertoires and intercomprehension skills

Intercomprehension refers to the practice of at least partially understanding a language that was never acquired previously by drawing on existing knowledge from other languages which are genealogically related to the target language (Doyé 2005). With regard to a potential multilingual benefit, Berthele (2011) provides evidence supporting the claim that multilinguals are on average better in developing receptive competences in previously unlearned languages compared to monolinguals. The observed advantage of multilinguals is not only explained by a higher amount of possible lexical transfer bases owing to more languages in the individual repertoires, but also by "a greater degree of flexibility in dealing with linguistic input that deviates from their own *L1* (*L2*, …, *Ln*) norm" (Vanhove and Berthele 2017: 25). The degree of formal deviation of the target structure from its counterparts in the source languages thus impedes less on the inferencing process in multilingual speakers and leads to better results in word recognition. Apart from the number of known languages, the level of proficiency and the relations of languages present in the individual repertoire also turned out to be significant predictors of correct cognate

guessing: Multilinguals with high proficiency in two languages (including dialects) that are closely related to the target language achieved the highest scores among all groups investigated in Berthele (2011). Other significant participant-related contributors to successful recognition of word meaning in cognates included age at testing and vocabulary learning ability. Structural constraints, such as formal distance between the target items and their cognates in the source languages, also played a significant role regarding the accuracy in inferring the correct word meaning. However, the formal distance effect seems to be systematically dependent on participant-related factors such as the individual levels of linguistic experience and cognitive capacities ("fluid intelligence") (Vanhove and Berthele 2017).

3 Study design

3.1 Participants

For our study we gathered data from thirty four adolescent speakers of a Slavic heritage language (HL) living in Germany. Nineteen participants were raised in a family where at least one parent used Russian when speaking to the child; the other fifteen grew up in Polish-speaking families. All participants were on average fifteen years old when they took the first intercomprehension test (see Table 1). They exhibit the typical profile of a heritage speaker (HS), i.e. they were either already born in Germany (twenty two participants) or came to Germany prior to the onset of schooling. With regard to their educational background, the sample is rather homogeneous: twenty six participants (75%) attended higher tracks of secondary school education at the time of testing, seven went to junior high schools, while one participant came from a vocational training school. A very distinctive feature of our sample is the high ratio of participants who received at least some formal instruction in their Slavic HL: Only eight participants never attended HL classes. While the extent of received HL instruction varies between our participants (see Brehmer and Mehlhorn 2018), almost 75% attended the respective courses for three years or more. Consequently, they show an overall high proficiency in their HL. This is reflected in the results obtained from an extensive test battery on HL proficiency that all participants had to take. Data on their proficiency in German were also available.[1] For

---

[1] The data come from a longitudinal project on the linguistic development of adolescent HSs of Russian and Polish in Germany (funded by the German Ministry of Education and Research, grant no. 01JM1302/1701). During the first two waves of data collection (2014 and 2015), we used an extensive test battery targeting different kinds of proficiency in both the heritage and the majority language (see Brehmer and Mehlhorn 2015 for more details).

the purpose of the current analysis, we will use the scores obtained in a vocabulary translation task (see Brehmer et al. 2017 for more details) and in a cloze test on grammatical proficiency as a general proxy of their attainment in the HL (maximum score in both tasks: 100, see Table 1). In a German vocabulary test and C-Test, the participants performed on average at ceiling (maximum score in both tasks: 100, see Table 1) which attests their high attainment in the majority language.

The current study builds on a comparison of the bilingual HS group with a control group of monolingual German peers. To obtain a comparable reference sample, we asked our bilingual participants to recruit monolingual controls themselves by inviting their best friend among their classmates to join the intercomprehension tasks. The only prerequisite was that the peers must have been raised in German only. As a result, we received a control group of thirty three monolingual German adolescents[2] who are comparable to our focus group in terms of age and educational background (see Table 1). All participants completed a brief questionnaire in order to gain background information on previously acquired foreign languages. They were also asked to deliver their grades obtained in foreign language subjects. As our participants mostly attend higher tracks of secondary education, almost all of them have been acquiring at least two foreign languages at school: English is the first foreign language for all participants. Only five bilingual and three monolingual participants acquire solely English at school, the rest of the sample is learning at least one additional language (mostly French, Spanish or Latin). We used the school grades obtained by the participants in their first foreign language English as an indicator of their foreign language learning aptitude. The grades range from 1 (best) to 6 (worst). Table 1 provides an overview on the demographic and language-proficiency variables of the bilingual and monolingual group:

Table 1: Demographic and language-proficiency variables.

Variable	Bilinguals		Monolinguals	
	Mean	**SD**	**Mean**	**SD**
Female/male (*n*)	15/19		21/12	
Age (years)	15.0	0.78	15.3	1.01
German vocabulary score	97.7	3.11	–	
German C-Test score	89.1	9.54	–	
Heritage language vocabulary score	80.5	12.17	–	
Heritage language grammar score	75.2	19.54	–	
English grade	2.5	0.90	2.5	0.85
Number of languages	4.0	0.69	3.0	0.52

2 One HS failed to recruit a monolingual peer.

The data reveal that we are dealing with comparable groups regarding school type, age and English language proficiency. The bilingual children differ from their monolingual peers in their bilingual upbringing, which results in one additional language (= HL) in their linguistic repertoire.

3.2 Materials and methods

For data gathering, we administered two paper-and-pencil intercomprehension tasks to our participants: First, both the bilingual and the monolingual participants received a text about the famous British singer Adele which was taken from the Swedish version of the online encyclopedia Wikipedia. For testing, we used the paragraphs that contained some biographical information about the artist and main facts about her career (albums and songs, awards). Owing to the text type (encyclopedia entry) and content, the text showed a high proportion of dates and English proper names, e.g. names of songs and albums (*Skyfall, Hello, Rolling in the deep*), films (*Bond*), awards and award categories (*Grammy, Oscar, Best Female Pop Vocal Performance*) and places (*London, Tottenham*). Furthermore, it contained a lot of borrowings from other languages, mostly from English (e.g. *singel* 'single', *radiohit, soul, jazz, film*), but also international lexical items with Greek or Latin roots (e.g. *november, gala, album*). We expected that this would enhance text comprehension and, at the same time, act as a stimulus for the participants to resort to their full linguistic repertoire when solving the task.

After exposing our informants to a text written in an unknown Germanic language, the second trial targeted an unknown Slavic language (Serbian). For reasons of comparability, we again selected a text about a famous pop artist (this time the American singer, songwriter and actor Justin Timberlake) taken from the Serbian version of Wikipedia.[3] As with the first text, we stuck to the paragraphs on Timberlake's biography and career. Again, it contained a lot of (English) song titles (*Like I love you, Cry me a river, Justified*), proper names (*N'Sync, Britni Spirs, Kameron Dijaz, Superboul*), and borrowings from other languages (*producent* 'producer', *solo album, singl* 'single', *bend* 'band').

In contrast to the Swedish text, there was no common source language for all groups of participants. Whereas in the Swedish test condition all participants could equally resort to either German or English as common Germanic source languages to extract the main information, the HSs would have been clearly in a better position than the monolingual controls in solving the Serbian task, as all of them have a Slavic

---

3 As Serbian uses both the Roman and the Cyrillic script, we decided to use the version with Roman characters in order to make the task feasible for all participants.

source language (either Russian or Polish). For most of the monolingual controls,[4] apart from general knowledge about Justin Timberlake, only accidental English or German cognates could support text decoding in the Serbian task. Although some scholars underline the possibility of applying intercomprehension tasks also with regard to languages from different language families (e.g. Morkötter 2019: 324), we finally refrained from administering the Serbian text to our monolingual controls. For this reason, the Serbian task only serves as a testing case to examine whether the HSs perform better in the Swedish or Serbian intercomprehension task or whether they do equally well in both tasks. In the latter case, this could be taken as a strong hint that the Slavic heritage language is fully accessible to them as a transfer base.[5]

In both trials, the participants were first asked to solve the tasks individually. The trials were conducted in a quiet place in the participants' homes[6] with one test instructor being constantly present to prevent the participants from using additional tools like online dictionaries of the target languages etc. The participants had to read the text carefully and try to understand its main content by drawing on their whole linguistic repertoire of previously acquired languages. There was no preceding training session. Thus, we are dealing with spontaneous or natural intercomprehension (Bär 2011: 28–29) in our study where the test-takers are not trained in relevant (meta) cognitive strategies on how to decode a text in an unknown language that is genealogically related to languages they already know. For each text, the participants had to solve three tasks: First, they had to assign ten lexical items from the text to parts of speech (POS). To exclude the possibility that a lack of grammatical terminology prevents our informants from coping with this task, we offered examples from German for each relevant POS. Secondly, they had to translate the selected items into German. In the third task, the test-takers were asked to write down all pieces of information they could extract from the text. For every fully provided piece of information we granted one point (maximum score: eighteen points/text). The participants could solve the individual tasks without any time pressure.

Due to the fact that the items from task one and two were embedded in an authentic text, the participants had all contextual information available in order to decide on the POS and the meaning of the respective items. This allowed us to test not only cognates, as had been done in many previous studies (see Section 3), but also opaque items, i.e. items which show only minimal or no formal overlap in the source and target languages (Klein and Stegmann 2000). Hypotheses regarding the meaning

---

4 There were only three monolingual peers who attended classes in Russian as a foreign language.
5 As Serbian belongs to the South-Slavic subgroup, neither the Russian (Russian being an East Slavic language) nor the Polish group (Polish belonging to the West Slavic branch) have a HL that is *per se* genealogically closer to Serbian.
6 The monolingual controls were invited to the home of their peers, but tested in a separate room.

and grammatical structure of these items can thus be data-driven, i.e. they are informed by formal overlaps with items in the source languages, or concept-driven, i.e. guided by the inferred content of the passage where the item occurs. Whether one or the other strategy is used to decode individual items, depends on the level of proficiency and the type and number of source languages available to the test-taker, the degree of interconnectedness of the lexical items in the multilingual mental lexicon, or the degree of familiarity of the test-taker with the content that is (supposedly) reported in the text. Both strategies can be applied conjointly or alternatively, depending on the individual item (Müller-Lancé 2006). In order to verify which of these strategies were applied by our participants in solving the tasks, we introduced an additional step in the data collection procedure: For the Swedish text, we asked the heritage bilingual and his/her monolingual friend to discuss their solutions together after they had completed the tasks individually. They were thus encouraged to (conjointly) produce think-aloud protocols in which they should explicitly describe their line of reasoning and the (meta)linguistic hypotheses that guided their individual guesses. The verbal protocols were recorded and later transcribed.

Previous research revealed that the degree of formal overlap has a significant impact on the successful establishment of interlingual correspondences between lexical items in the target and source languages (see Section 2). For this reason, the selection of individual test items from the Swedish text was guided by two criteria: (a) every main POS should be represented in the sample; (b) the items should exhibit different degrees of formal overlap. We therefore divided the items into several classes:

Class 1a: target items that show cognates in all involved languages, e.g. Swedish *brittisk*, English *British*, German *britisch*, Russian *britanskij*, Polish *brytyjski* ($n = 3$);

Class 1b: target items that exhibit a transparent morphemic structure which allows to establish interlingual correspondences at least for some morphemic components in all languages involved, e.g. Swedish *artistnamn* 'pseudonym, screen name', cf. English *artist* + *name*, German *Artist* 'performer' + *Name* 'name', Russian *artist*, Polish *artysta* ($n = 1$);

Class 2: target items that have cognates in both Germanic, but not in the Slavic source languages, e.g. Swedish *i*, English and German *in* versus Russian *v*, Polish *w* ($n = 2$);[7]

Class 3: opaque target items that have no cognate in any of the source languages, e.g. Swedish *och* versus English *and*, German *und* and Russian/Polish *i* ($n = 4$).

---

7 Note that in this case there is even the potential for a distracting effect for participants from the bilingual group, as there is a homophonic conjunction *i* 'and' in Russian and Polish.

It goes without saying that the formal similarities themselves vary between the cognates in the individual languages (cf. the equivalents of the Swedish item *brittisk* mentioned above). This obviously has an effect on the likelihood of successful interlingual identification by our participants (cf. Vanhove and Berthele 2017). Given the low number of test items and the diverse linguistic backgrounds of our participants, however, we did not measure the exact degree of formal discrepancy between the test items in the target and all source languages,[8] but restricted ourselves to the rather coarse classification described above.

For the Serbian task, the items were selected again in order to represent different POS. Furthermore, we chose items that

Class 1a: have cognates in both Slavic HLs, e.g. Serbian *učestvuje* '[s/he] participates', Russian *učastvuet*, Polish *uczestniczy* (*n* = 6);

Class 1b: exhibit a transparent morphemic structure which allows to establish interlingual correspondences for some morphemic components in both HLs, e.g. Serbian *tekstopisac* 'songwriter', cf. Russian *tekst* + *pisat'* 'to write', Polish *tekst* + *pisać* 'to write' (*n* = 1);

Class 2: have cognates in only one Slavic HL, e.g. Serbian *godina* 'year', Russian *god* versus Polish *rok* (*n* = 2);

Class 3: represent opaque items, e.g. Serbian *više* 'more (than)', Russian *bol'še*, Polish *więcej* (*n* = 1).

3.3 Research questions

The main objective of the current paper is to explore whether HSs make use of their HL when they are put in a situation where they have to perform a multilingual search across their lexicon. Furthermore, we want to contribute to the ongoing discussion of whether bilinguals have a general cognitive and/or linguistic advantage over monolinguals that contributes to a better performance in tasks on receptive multilingualism. Thus, our main research questions are the following:

RQ 1: Do the HSs outperform the monolingual controls in an intercomprehension task with Swedish as the unknown target language?

RQ 2: Do HSs perform better in an intercomprehension task with a Germanic target language (Swedish) than with a Slavic target language (Serbian)?

---

8 Formal distance between cognates in different languages is normally measured by means of the Levenshtein algorithm which is based on the number of insertions, deletions or substitutions necessary to transform an orthographic string of a given target item into its equivalent item in the source language (see Vanhove and Berthele 2017: 28, with additional references on the impact of Levenshtein distances on interlingual identification).

RQ 3: Which strategies do HSs apply when decoding items and text passages in an unknown target language and which role does the Slavic HL play in these strategies?

RQ 4: Does the existence of cognates facilitate successful comprehension of unknown target items and the identification of their grammatical structure?

Although the role of person-related factors for successful interlingual inferencing is not a central issue in our analysis, two factors that turned out to be relevant in previous research will be looked at: (i) breadth of the individual linguistic repertoire, and (ii) level of proficiency in English and the Slavic HL.

4 Results

4.1 Overall performance in the Swedish trial

We are first turning to the description of the results regarding the performance of both groups in the Swedish trial. Figure 1 depicts the average relative correctness scores in the three subtasks.

Both groups performed equally well in all three subtasks. A t-test for independent samples revealed no statistically significant difference between the bilinguals and monolinguals either for determination of POS ($t(65) = 0.08$, $p = 0.937$) or for overall text comprehension ($t(65) = -0.71$, $p = 0.482$), and only a weak significance
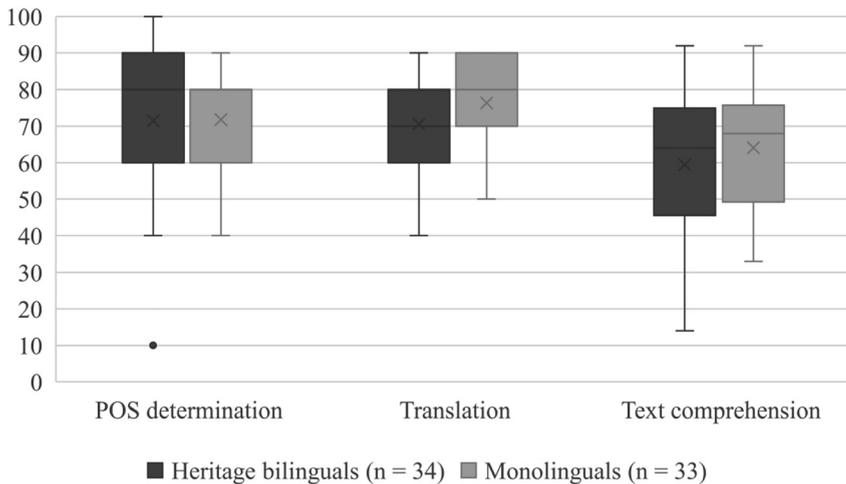

Figure 1: Relative correctness scores (in %) for subtasks in the Swedish trial.

for translation ($t(65) = 1.73$, $p = 0.089$). Effect sizes are very small (Cohen's d for POS determination = 0.019; for translation = 0.422; for overall text comprehension = 0.226). On average, the bilinguals scored 7.1 points (out of a maximum of ten) (SD = 1.94) and the monolinguals 7.2 (SD = 1.61) in the task on POS determination. When asked to translate the ten selected Swedish items, the HSs scored again on average 7.1 points (SD = 1.32) and the monolinguals 7.6 (SD = 1.41). It might come as a surprise that for some individuals the scores for POS determination are lower than for translating the same items. Normally one would expect an implicational relationship between the results in these two tasks,[9] i.e. the selection of a translational equivalent for the Swedish item should depend to a certain extent on the decision regarding the assumed POS that this item represents. However, our informants sometimes failed to determine the POS, but nevertheless provided correct translations for the same items. We interpret this 'mismatch' as a consequence of general problems in determining POS, irrespective of the language, despite our effort to offer prototypical examples for each POS in German at the beginning of the task. For example, some participants considered the Swedish preposition *i* 'in' to be a conjunction, but translated it correctly by using the German equivalent preposition *in*. Likewise, some informants believed the Swedish adjectives *brittisk* 'British' and *internationell* 'international' to be nouns, but they provided the German adjectives *britisch* and *international* as correct translation equivalents. Interestingly, though, the reverse mismatch occurred very infrequently in our data, e.g. when the Swedish pronoun *hon* 'she' had been correctly identified as a pronoun in the task on POS, but was translated by the same participant as the German preposition *für* 'for'. In general, the task on POS determination obviously presupposes a high degree of explicit metalinguistic knowledge that is sometimes lacking (maybe due to a lack of attention in school regarding formal properties of words), while the search for a functional equivalent in translating the same items implicitly led to the correct decision concerning the required POS. The results obtained in the POS determination and translation tasks showed only a moderate positive correlation for both groups (bilinguals: $r = 0.63$, $p < 0.001$, monolinguals: 0.54, $p = 0.001$). The task to extract as many pieces of information as possible from the entire text turned out to be more difficult for both monolingual and bilingual participants: On average the bilinguals scored 10.7 points (out of a maximum of eighteen) (SD = 3.86), while the monolinguals performed slightly better (average score: 11.5, SD = 3.30). Translation of individual items and overall text comprehension were again moderately correlated in both groups (bilinguals: $r = 0.55$, $p < 0.001$, monolinguals: 0.60, $p = 0.018$).

---

9 We are very much indebted to one of the anonymous reviewers for drawing our attention to this fact.

4.2 Comparison of performance in the Swedish and Serbian trial for the bilingual group

As the heritage bilinguals completed both intercomprehension tasks, we can compare the effect of the language family of the target language (Germanic vs. Slavic) on the performance in the three subtasks. Although the two texts had been selected in such a way that they closely resembled one another with regard to content, text type and degree of syntactic complexity, they still represent different texts with different languages and different lexical items. Therefore, the results cannot be compared in a straightforward manner, which is why we refrained from running a statistical analysis regarding the significance of the differences that we found in the results obtained from these two tasks. Nevertheless, a descriptive account of the performance of the heritage speakers in both tasks already reveals interesting tendencies. Figure 2 shows the results of the HSs in the Swedish and Serbian trials.

The HSs exhibited a considerably weaker performance in all Serbian trials compared to the Swedish ones. The differences emerge for all three subtasks: In the Serbian trials, the average score is between one and four points lower compared to the Swedish trials, depending on the task: POS determination: 5.9 (SD = 1.73), translation: 5.2 (SD = 1.72), overall text comprehension: 6.4 (SD = 2.82). As in the Swedish trials, the results in the Serbian POS determination and translation tasks showed a moderate positive correlation ($r = 0.68$, $p < 0.001$). The same applies to
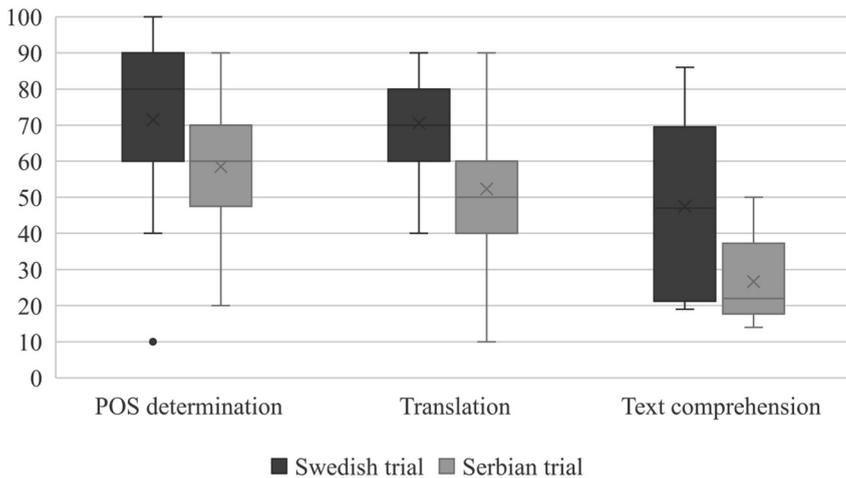
Figure 2: Relative correctness scores (in %) for subtasks in the Swedish and Serbian trials (*n* = 34 bilingual participants per trial).

the results from the Serbian translation and text comprehension tasks ($r = 0.58$, $p < 0.001$). If the Swedish and Serbian tasks are compared to one another, we find a moderate positive correlation between the results obtained in both POS determination ($r = 0.52$, $p = 0.001$) and translation ($r = 0.43$, $p = 0.011$) tasks. We found the strongest positive correlation in our whole data set for the performance of the HSs in both tasks on text comprehension ($r = 0.82$, $p < 0.001$).

4.3 Performance with regard to individual items

4.3.1 Swedish items

In order to explore the effect of formal similarity between target and source language items on POS and word meaning recognition, we divided our set of ten selected items into several groups, depending on the existence of cognates in the Germanic and/or Slavic source languages (see Section 3.2). As the numbers of items that comprise the different categories are too small to make any generalizable statements, our data can just offer tentative results regarding the impact of formal similarity on word recognition. Table 2 shows the ratio of successful POS and word meaning recognition in the different groups of Swedish items.

The availability of at least one Germanic cognate seems to enhance the possibility of correctly assigning an unknown Swedish item to a POS for both groups. This holds for nouns (*artistnamn* 'screen name'), adjectives (*internationell* 'international'), numerals (*första* 'first'), verbs (*nominerad* 'nominated'), and prepositions (*i* 'in'). The additional existence of a cognate in the Slavic HL did not lead to higher correctness scores in the bilingual group (cf. performance for items of class 1a and 2). Opaque items posed the biggest problems for both groups. This might be related to a high proportion of function words in this group. Especially the Swedish article *ett* yielded very few instances of correct POS determination in both groups,

Table 2: Degree of formal transparency of Swedish items and ratio of successful POS and word meaning recognition (in %) in bilinguals ($n = 34$) and monolinguals ($n = 33$).

Item class	POS recognition		Word meaning recognition	
	Bilinguals	**Monolinguals**	**Bilinguals**	**Monolinguals**
1a ($n = 3$)	83.3%	86.9%	97.1%	97.0%
1b ($n = 1$)	91.2%	97.0%	58.8%	81.8%
2 ($n = 2$)	83.8%	86.4%	94.1%	95.5%
3 ($n = 4$)	51.5%	47.7%	44.1%	50.0%

despite the presence of articles in German and English. The possessive pronoun *henne* 'her' and the personal pronoun *hon* 'she' were correctly identified only by half of the bilingual and monolingual participants, while the conjunction *och* 'and' was the easiest item in this class, again for both groups. Word meaning recognition was especially successful for items belonging to class 1a and 2. Bilingual speakers performed on average slightly worse in inferring the meaning of function words from the given context than their monolingual peers. In particular, the Polish group had problems in translating the conjunction *och* 'and' (which was often rendered as German *auch*, possibly due to the influence of the homophonic German dialectal variant [o:x] of this conjunction) and the personal pronoun *hon* 'she'. In the latter case, this was linked to a poorer performance in the POS determination task compared to the Russian group. The Russian group was far more successful in identifying *hon* as a personal pronoun (79 vs. 47% correct guesses). This interpretation might have benefitted from an interlingual correspondence with the Russian personal pronouns *on* 'he' and *ona* 'she'. The same correspondence would have been available for the Polish group (where *on* and *ona* also function as third person pronouns), but for some reason most of the Polish HSs failed to establish such an interlingual correspondence or they were confused by the fact that *hon* in the Swedish text obviously referred to a female singer. The poorer performance of the bilingual group in translating the item *artistnamn* 'screen name' was mainly due to the fact that the bilinguals more often neglected the second component of the compound in their translations, thus dropping *-name* from the correct German equivalent *Künstlername*. This might be linked to a lower number of compounds in Slavic languages compared to German.

4.3.2 Serbian items

The Serbian lexical items were also divided into three subclasses depending on the (non-) existence of cognates in the two HLs (see Section 3.2). Table 3 offers an overview on the ratio of successfully determined POS and word meanings in the bilingual group.

Table 3: Degree of formal transparency of Serbian items and ratio of successful POS and word meaning recognition (in %) in bilingual speakers ($n = 34$).

Item class	POS recognition	Word meaning recognition
1a ($n = 6$)	59.8%	52.9%
1b ($n = 1$)	85.3%	82.4%
2 ($n = 2$)	63.2%	52.9%
3 ($n = 1$)	14.7%	17.6%

At first glance, the existence of cognates in both HLs (class 1a) seems to not necessarily facilitate the decoding of an unknown Serbian lexical item, as regards the POS as well as its meaning. Despite the structural overlap of the three Slavic languages with regard to basic grammatical categories, language-specific peculiarities regarding the inflectional markings of content words and diachronic sound changes obviously impede the establishment of interlingual correspondences. Thus, the bilingual participants succeeded in only approximately half of the cases in assigning a Serbian item to the correct POS and in decoding its meaning, although there would have been transfer bases available in both HLs (class 1a). The closer the formal similarities between the target and the source items of this class, the more accessible were the items to our participants: Thus, the Serbian reflexive possessive pronoun *svoj* was recognized by almost all participants (91.2%) and subsequently correctly translated (97.1%) due to its formal closeness to the Russian (*svoj*) and Polish (*swój*) counterparts. In case of formal discrepancies, the degree of recoverability obviously depended on the context: While the Serbian preposition *u* 'in' was correctly identified as a preposition by two thirds of our bilinguals (64.7%), the diachronic sound change /v/ to /u/ in Serbian (cf. its Russian and Polish counterpart *v/w*) hampered the correct decoding of its meaning (only 17.6% correct guesses). The relevance of formal distance is also revealed by the items of class 2: Here, speakers of the HL, where a cognate existed, performed better than the bilinguals whose HL lacked a morphosemantically similar item. While the Serbian superlative *najveći* 'biggest' was correctly identified as an adjective by more than half of the Russian (53%) and Polish (60%) HSs due to the common Slavic superlative prefix {/naj/}, the Russian participants performed worse in decoding its meaning (16%) than their Polish-speaking peers (67%) who seemingly benefitted from the formal similarity of its Polish equivalent *największy* (cf. Russian *naibol'šij*). Opaque items (class 3) turned out to be the most problematic for our bilinguals, as had been expected.

4.4 Strategies applied in interlingual inferencing

The data from the thinking-aloud task shed light on the strategies that were applied by our participants in order to comprehend the Swedish text and to decode the individual Swedish items. They used both form- and content-driven approaches as mentioned in Section 3.2. For our purpose, the role of the individual linguistic repertoires in solving the tasks is especially relevant. The transcripts confirmed that the participants systematically resorted to previously acquired languages to establish interlingual correspondences and, by doing so, to be able to infer the meaning of individual Swedish items and/or whole text passages. As expected, German and English figure as most often mentioned sources for drawing inferences regarding form and meaning of Swedish items. Formal (graphemic) similarities are the main

triggers for establishing interlingual correspondences. This was observed for Swedish items like *ledmotiv* < German *Leitmotiv* 'leitmotif', *belönad* < German *belohnt* 'rewarded', *fyra* < German *vier* 'four', *handlar* < German *handeln* 'to be about sth.', *första* < English *first* or *kallat* < English *called*. However, the bilinguals also occasionally referred to their Slavic HL as an additional clue to substantiate their guesses. This was the case with Swedish numerals like *tredje* 'third' < Russian *tri* 'three' and *tretij* 'third' or *två* < Russian *dva* 'two'. Furthermore, the first component of the Swedish compound *artistnamn* was also identified by two Russian HSs by drawing on Russian *artist* 'artist'. One of them also offered lexical parallels from English (*artist*) and French (*artiste*) to substantiate her claim that the correct German translation should be *Künstler*. In one instance, the correct translation for *i* 'in' is derived from the co- and context, while an alternative form-driven approach by using the Polish homophone *i* 'and' is explicitly dismissed by a Polish HS.

4.5 Effects of person-related properties on performance in the tasks

Although an in-depth analysis of the effects of different participant-related variables on test performance lies outside the scope of the current study, we would like to add some observations on correlations between personal properties of the bilingual test-takers and observed results in the intercomprehension tasks. Among the factors we looked at were proficiency in the Slavic HL (vocabulary score and grammar score, see Table 1), proficiency level in English (measured by self-reported school grade in English at the time of testing), and number of learned languages. Correlation analyses revealed mostly weak correlations between the performance in the Swedish and Serbian intercomprehension tasks and the personal properties of the bilingual speakers. Moderate positive correlations could be established only for the vocabulary score in the Slavic HL and translation of Serbian items ($r = 0.63$, $p < 0.001$) and the number of learned languages and POS recognition in Serbian items ($r = 0.50$, $p = 0.002$). Furthermore, a relatively strong association was found for the grades obtained in English and the translation of Swedish ($r_s = 0.40$, $p = 0.019$) and Serbian items ($r_s = 0.42$, $p = 0.013$) as well as overall text comprehension in the Swedish trial ($r_s = 0.46$, $p = 0.007$).

5 Discussion

The results of the current study offer no evidence for a general bilingual advantage of heritage bilinguals over monolinguals in solving tasks on receptive multilingualism

that involve unknown Germanic languages. The bilinguals did not outperform their monolingual peers in the tasks of the Swedish trial which aimed at (i) assigning given Swedish items to POS, (ii) translating these items, and (iii) overall text comprehension. The monolingual controls even performed slightly better than the bilinguals, although the difference in the translation task yielded only marginal statistical significance ($p < 0.1$). One could argue that due to the fact that the target language Swedish belonged to the Germanic language family, a beneficial effect of having an additional Slavic language as a potential transfer base could not be expected. Proficiency levels in previously acquired Germanic languages, in our case German as the $L1$ or early acquired $L2$ of all participants, and English as the first foreign language for all test-takers, seem to matter more than having a second $L1$ which genealogically belongs to another Indoeuropean language family (cf. also Berthele 2011). As our heritage bilinguals exhibit a high proficiency level in German and perform on average on par with the monolingual peers in English as a foreign language (see Table 1), the observed "null effect" might not come across as very surprising. However, the bilingual advantage discussed in previous research is not always linked to linguistic knowledge (e.g. more lexical transfer bases), but also to conceptual knowledge and more general cognitive skills like executive control or higher metalinguistic awareness in bilinguals (see Section 1). Our findings are thus in line with previous research that casted doubt on a general bilingual advantage, at least in heritage speaker contexts (see Berthele and Udry 2019; Berthele and Vanhove 2020; Lorenz et al. 2020 for similar results).

Our results, however, provide evidence that at least for some individuals in our sample the HL is indeed used as a resource to decode items or text passages in an unknown Germanic language. Some of our HSs established interlingual correspondences between the Swedish target item and lexical units in their HL in order to infer its meaning or to assign it to POS (cf. the correct identification of Swedish *hon* as a third person pronoun by Russian HSs). Sometimes these supposed Slavic-Germanic interlingual correspondences were backed by adducing further evidence from additional languages (English, French), pointing to a high degree of interconnectedness of lexical units in the individual's multilingual mental lexicon (cf. Pavlenko 2009). The fact that we could observe this strategy only in very few bilingual participants of our sample shows, however, that HLs seem to not always be readily available as a resource for solving tasks on receptive multilingualism. The influence of the majority language (German) and the first acquired foreign language (English) on strategies used to infer the meaning of unknown target words is definitely stronger than positive interdependence effects stemming from the HL (cf. also Edele et al. 2018).

This claim is further corroborated by the results of the bilingual participants in the Serbian trial. When the Slavic HL turns into the primary (or most natural) transfer base for solving the tasks, the bilingual participants perform worse in all of the tasks

compared to the Swedish trial. The weaker performance in the Serbian trial might also be related to the fact that among the selected items were no loan words comparable to *internationell* 'international' or *nominerad* 'nominated' in the Swedish trial. Furthermore, there is no second source language available which is genealogically closely related to the target language. It has been shown that interlingual inferencing is more efficient if test-takers master two languages (or varieties) which are related to the unknown target language (Berthele 2011). This was the case for the Swedish trial (German, English), but not for the Serbian (only Russian or Polish). Still, at least with regard to individual lexical items, the HSs manage to assign them to appropriate POS and infer their meaning in more than half of the cases on average (see Section 4.2). The more similar the form of the target item to its equivalent in the Slavic HL, the better are the results in establishing interlingual correspondences that lead to correct guesses regarding POS or word meaning. But even if cognates are available in the HL, POS determination and item comprehension do not substantially exceed the level of recovery rates of opaque items in a Germanic target language. Given that our participants were never explicitly trained in establishing interlingual correspondences between their HL and other Slavic languages and are thus unfamiliar with the diachronic sound changes that brought about quite different phonological (and, subsequently, graphemic) representations of the same Slavic etymological units, the ratio of successful guesses in the Serbian trial is nevertheless quite impressive. While individual items which are contextually embedded seem to be quite accessible for our HSs, they often fail to extract essential pieces of information from the Serbian text, although it contains a lot of information (proper names, dates, English song titles) that should have been helpful in reconstructing the essence of the text by resorting to non-linguistic world knowledge. The strong correlation we found between the performance in the Swedish and Serbian tasks on overall text comprehension indicates that general cognitive abilities related to decoding texts in unknown languages (cf. Müller-Lancé 2006; Meißner 2011) seem to be essential factors for predicting test outcomes.

The degree of formal similarity between the items in the target and source languages turned out to be a likely candidate for improving the probability of a successful establishment of interlingual correspondences. In the Swedish trial, the existence of cognates in at least one of the Germanic source languages seemed to suffice to trigger interlingual inferencing. Additional transfer bases in the Slavic HLs did not lead to a higher ratio of correct guesses in the bilingual group. In the Serbian trial, POS recognition was more independent from interlingual distances between Serbian target and Russian/Polish source items. With regard to word meaning, almost or completely identical phonemic or graphemic representations of target and source items obviously facilitated the decoding of the meaning of the target item. The greater the differences regarding prefixes, suffixes, and/or inflectional endings of target and source items, the weaker was the performance in the task. Furthermore, our results seem to confirm previous findings (see, e.g., Berthele 2011; Möller and Zeevaert 2010)

that interlingual inferencing is especially efficient when the first segments of the items are identical, with identical consonantal patterns at the onset being more important than vowel representations, cf. higher correctness rates in word meaning retrieval of items like Swedish *första* < English *first*, Serbian *svoj* 'his/her' < Russian *svoj*, Polish *swój*, Serbian *učestvuje* '[s/he] participates' < Russian *učastvuet*, Polish *uczestniczy* compared to items like Serbian *nastup* 'gig' < Russian *vystuplenie*, Polish *występ*. However, apart from form-driven approaches, participants of both groups also frequently referred to concept-driven strategies where the POS or the meaning of the target item was established due to close consideration of the co- and context where it occurred (relationship to other lexical items in the immediate neighborhood, text type, non-linguistic information given in the co- and context etc.).

Regarding personal properties of the test-takers, our study could only confirm that proficiency in English positively impacts performance in the Swedish and, to a lesser extent, in the Serbian trial. For the Serbian translation task, vocabulary knowledge in the related HL also mattered (cf. Berthele 2011 or Vanhove and Berthele 2015 for similar results). Our sample consists of speakers who exhibit a rather high level of proficiency in their HLs which is also related to the fact that most of them had at least some access to formal instruction in it. Formal instruction in the HL seems to matter for a successful performance in intercomprehension tasks (see Brehmer and Arifulin to appear), even though a more balanced sample with regard to this factor would be needed to substantiate this claim. Formal training in the HL in all likelihood increases explicit knowledge and, subsequently, metalinguistic awareness regarding rules and regularities in the HL, which in turn could positively impact the ability to identify parallel structures in a related, but unknown language.

6 Conclusions

Our initial assumption that heritage bilinguals are particularly efficient in recognizing interlingual correspondences between items in one or more of their languages and an unknown language that is related either to the majority or the heritage language was not confirmed by the data. They performed on par with the monolingual controls in the Swedish intercomprehension task. Furthermore, the interlingual heuristics developed by the heritage bilinguals seemed to draw more from the Germanic components of their linguistic repertoire than from the Slavic HL, despite a relatively high proficiency in the latter. Thus, we failed to substantiate a bilingual advantage for our heritage bilinguals. This is in line with previous research that could also not observe a multilingual boost in heritage speaker contexts when relevant personal variables (like age, educational background, size and make-up of language repertoires, level of proficiency in *L*1 and foreign

languages like English) were controlled for, as in the current study. However, when put into a situation where the Slavic HL serves as the most suitable transfer base because it is genealogically closest to the unknown target language, the bilingual participants were able to develop interlingual heuristics that helped them to identify structures and meanings of unknown Serbian items. On average, however, interlingual inferencing worked more poorly in this context and was more confined to single items, especially cognates with a high degree of formal overlap. Participant-related factors (e.g. vocabulary knowledge in the HL) seem to have an impact on the degree to which cross-linguistic similarities are perceived. HSs are thus endowed with a potential for developing receptive multilingualism in languages closely related to their HL, but access to potential transfer bases is often blocked or restricted to very obvious cases with high degrees of formal similarity. Systematic instruction in the HL might provide a starting point to raise the awareness of interlingual correspondences involving, but not confined to, unknown languages genealogically related to the heritage HL.

References

Adesope, Olusola O., Tracy Lavin, Terri Thompson & Charles Ungerleider. 2010. A systematic review and meta-analysis of the cognitive correlates of bilingualism. *Review of Educational Research* 80(2). 207–245.

Bär, Marcus. 2011. Interkomprehension – Nicht ‚nur' ein Weg zur Förderung individueller Mehrsprachigkeit. *Die Neueren Sprachen* 2. 21–32.

Berthele, Raphael. 2011. On abduction in receptive multilingualism: Evidence from cognate guessing tasks. *Applied Linguistics Review* 2. 191–220.

Berthele, Raphael & Isabelle Udry. 2019. Multilingual boost versus cognitive abilities: Testing two theories of multilingual language learning in a primary school context. *International Journal of Multilingualism*. 1–20. https://doi.org/10.1080/14790718.2019.1632315.

Berthele, Raphael & Jan Vanhove. 2020. What would disprove interdependence? Lessons learned from a study on biliteracy in Portuguese heritage language speakers in Switzerland. *International Journal of Bilingual Education and Bilingualism* 23(5). 550–566.

Bialystok, Ellen. 2001. *Bilingualism in development: Language, literacy, and cognition*. New York: Cambridge University Press.

Brehmer, Bernhard & Grit Mehlhorn. 2015. Russisch als Herkunftssprache in Deutschland: Ein holistischer Ansatz zur Erforschung des Potenzials von Herkunftssprachen. *Zeitschrift für Fremdsprachenforschung* 26(1). 83–121.

Brehmer, Bernhard & Grit Mehlhorn. 2018. Unterricht in den Herkunftssprachen Russisch und Polnisch: Einstellungen und Effekte. In Grit Mehlhorn & Bernhard Brehmer (eds.), *Potenziale von Herkunftssprachen: Sprachliche und außersprachliche Einflussfaktoren*, 259–292. Tübingen: Stauffenburg.

Brehmer, Bernhard, Tatjana, Kurbangulova & Martin Winski. 2017. Measuring lexical proficiency in Slavic heritage languages: A comparison of different experimental approaches. In Tanja

Anstatt, Christina Clasmeier & Anja Gattnar (eds.), *Slavic languages in psycholinguistics. Chances and challenges for empirical and experimental research*, 225–256. Tübingen: Narr.

Brehmer, Bernhard & Vladimir Arifulin. to appear. „Wir understand svenska i srpski!" – Spontane Interkomprehension bei Herkunftssprecher*innen des Russischen in Deutschland. In Anastasija Kostiučenko, Tamara Münzer & Agnieszka Zawadzka (eds.), *Slawische Sprachen unterrichten*. Frankfurt/Main: Lang.

Cenoz, Jasone. 2003. The additive effect of bilingualism on third language acquisition: A review. *International Journal of Bilingualism* 7(1). 71–87.

Doyé, Peter. 2005. *Intercomprehension. Guide for the development of language education policies in Europe*. Strasbourg: Council of Europe.

Edele, Aileen, Sebastian Kempert & Kristin Schotte. 2018. Does competent bilingualism entail advantages for the third language learning of immigrant students? *Learning and Instruction* 58. 232–244.

Gathercole, Virginia C. M., Enlli M. Thomas, Ivan Kennedy, Cynog Prys, Nia Young, Nestor Viñas Guasch, Emily J. Roberts, Emma K. Hughes & Leah Jones. 2014. Does language dominance affect cognitive performance in bilinguals? Lifespan evidence from preschoolers through older adults on card sorting, Simon, and metalinguistic tasks. *Frontiers in Psychology* 5. 11.

Jessner, Ulrike. 2008. A DST model of multilingualism and the role of metalinguistic awareness. *The Modern Language Journal* 2. 270–283.

Klein, Horst G. & Tilbert D. Stegmann (eds.). 2000. *EuroComRom – Die sieben Siebe: Romanische Sprachen sofort lesen können*. Aachen: Shaker.

Lorenz, Eliane, Sharareh Rahbari, Ulrike Schackow & Peter Siemund. 2020. Does bilingualism correlate with or predict higher proficiency in L3 English? A contrastive study of monolingual and bilingual learners. *Journal of Multilingual Theories and Practices* 1(2). 185–217.

Meißner, Franz-Joseph. 2011. Teaching and learning intercomprehension: A way to plurilingualism and learner autonomy. In Inez de Florio-Hansen (ed.), *Towards multilingualism and the inclusion of cultural diversity*, 37–58. Kassel: Kassel University Press.

Möller, Robert & Ludger Zeevaert. 2010. "Da denke ich spontan an Tafel": Zur Worterkennung in verwandten germanischen Sprachen. *Zeitschrift für Fremdsprachenforschung* 21. 217–248.

Morkötter, Steffi. 2019. Interkomprehension und Transfer. In Christiane Fäcke & Franz-Joseph Meißner (eds.), *Handbuch Mehrsprachigkeits- und Mehrkulturalitätsdidaktik*, 321–324. Narr: Tübingen. https://doi.org/10.1515/9783110629378-021.

Müller-Lancé, Johannes. 2006. *Der Wortschatz romanischer Sprachen im Tertiärsprachenerwerb. Lernerstrategien am Beispiel des Spanischen, Italienischen und Katalanischen 2*. Aufl. Tübingen: Stauffenburg.

Paap, Kenneth R., Regina Anders-Jefferson, Lauren Mason, Katerinne Alvarado & Brandon Zimiga. 2018. Bilingual advantages in inhibition or selective attention: More challenges. *Frontiers in Psychology* 9. 1409.

Pavlenko, Aneta. 2009. *The bilingual mental lexicon: Interdisciplinary approaches*. Buffalo: Multilingual Matters.

Vanhove, Jan & Raphael Berthele. 2015. The lifespan development of cognate guessing skills in an unknown related language. *International Review of Applied Linguistics in Language Teaching* 53(1). 1–38.

Vanhove, Jan & Raphael Berthele. 2017. Interactions between formal distance and participant-related variables in receptive multilingualism. *International Review of Applied Linguistics in Language Teaching* 55(1). 23–40.