Processing of Written Sentences on the Example of Russian-English Bilinguals

LIUBOV DARZHINOVA

Abstract

Written language processing is broadly explained from the perspective of English and alongside with the other world languages. The current study views the concerned type of processing through the Russian-English linguistic combination, which is seen as having a limited research in the investigation into written language processing. By using a mixed methods research design, 21 Russian-speaking users of English were engaged in the study. The subjects filled in the questionnaire about their linguistic background followed by the participation in the experiment. The experimental setting was as follows: each subject had to classify the 240 Russian-English sentences as correct or not during the time interval of 5000 milliseconds. The results of the study conform with the BIA model, and suggest that semantic type of written language processing is more accurate than syntactic type but at around same level in time constraints.

Keywords
Bilinguals, English, language processing, Russian, written language, written sentences

Center for Research on Linguistics and Languages Studies (CRLLS), The Education University of Hong Kong, Hong Kong; liubov@s.eduhk.hk

Introduction

Processing of written language encompasses diverse stages of intricate cognitive processes (Kendeou et al., 2016). Consequently, syntactic information facilitates in combining the words we read in a way so that we can get the meaning of sentences which is semantics. To comprehend a sentence, a language user is supposed to perform a number of particular tasks concurrently, associating words with their meanings, which is semantics, and transferring or deriving meaning from the relationship among words in a sentence, which is syntax (Wu et al., 2016). Overall, both syntax and semantics add up to their equal functioning in language processing. In bilinguals, the situation is complicated by the need to choose a language for decoding. In other words, there is a competition of languages and the bilinguals' job is to correctly choose a communicative code among their language systems, while automaticity of perception between the signified and the signifier provides a fully-fledged perception and targeted information transfer (Abisheva, 2001).

The focus of applied linguistics on written language processing has lately been the research on individuals using behavioral methods among which are as error rates, reaction time measures and others. It is also found out that theoretical and experimental progress has been achieved by defining the semantic and syntactic types of operations that underlie language processing in English comparing with other world languages. It is observed the existing linguistic literature has a limited body of knowledge providing implications to the semantic and syntactic types of bilingual language processing. Russian, which is the eighth spoken language (Eberhard et el., 2019), still does not have an extensive research on this matter. However, to the best of the author's knowledge, the studies encompassing the Russian-English linguistic pair are up to now related to the issues of translation ambiguity (e.g., Jouravlev & Jared, 2019), cross-language cognates (e.g., Sherkina-Lieber, 2004), auditory language processing (e.g., Marian & Spivey, 2003), the relationship among syntactic, working memory, and phonology (e.g., Abu-Rabia, 2001), etc. To address the revealed gap, the current paper aims to address a long-standing linguistic dilemma by exploring semantic and syntactic aspects of written language processing based on the Russian-English linguistic combination. The study attempts to answer these two research questions:
1. Which type of written language processing on the sentence level in L1 and L2 is faster, semantic or syntactic?
2. Which type of written language processing on the sentence level in L1 and L2 is more accurate, semantic or syntactic?

Literature Review

Written language processing

Linguistics increasingly acknowledged the awareness of written language processing as a 'multisensory experience' to advance a deeper understanding of the events described (De Koning & van der Schoot, 2013). Readers need to create a consistent representation of the meaning by integrating the semantic properties of word by word in accordance with

certain syntactic rules (Hagoort, 2013). To reach a comprehension of what is written, readers need to integrate the situational information with their basic knowledge, by which a non-linguistic, comprehensible and cohesive mental picture of the 'state of affairs' described in the sentence will be constructed (Bråten et al., 2011).

Other studies suggest that processing of written language includes the re-activation of sensory, motor, as well as emotive experiences that are kept in areas of the brain responsible for real perception, action, and emotions, which the reader has attained throughout prior contacts in the real world (Barsalou, 2008; Pulvermüller, 2005; Zwaan & Radvansky, 1998). In particular, in accordance with the embodied concepts of cognition, readers build a mental simulation of the events taken place in the sentence (Kintsch, 1988; van den Broek, 2010). For example, to understand a sentence, such as 'She saw the egg in the skillet', it is necessary to reactivate the perceived information to imitate the shape of the object (an egg with the sunny side up), which is disguised in the sentence (Engelen, Bouwmeester, de Bruin, & Zwaan, 2011; Zwaan & Pecher, 2012). Pursuant these views, it may be reached that readers quickly integrate various types of information; thus, they decode lexical elements, create syntactic structures, and gradually assign the interpretation to the incoming line of words.

Written language processing in bilinguals

The comprehension of syntactic structure is of paramount importance in written language processing and is considered as the process of analyzing and determining the structure of a text which is made up of sequence of tokens with respect to a given formal grammar. It is also considered as a vital stage for semantic analysis (Mallamma et al., 2014). In the situation of bilingualism, language users deal with information from the other language, which becomes activated as well (De Groot, Delmaar, & Lupker, 2000).

Since typical literate native speakers easily complete these processes, it is suggested that the understanding of sentences in a second language (L2) is less effective. Processing of L2 sentences shows that building a syntactic structure is more difficult than understanding of sentences in L1 (Roberts, 2013). Differences in native and non-native syntactic processing may reflect quantitative delays in lexical access that are related to L2 processing. These slowdowns can have the 'hook' effect in that the construction of the syntactic structure is delayed or weakened compared to native processing (Hopp, 2015). Based on that, it can be presumed that due to the complexity of linguistic phenomena across the world languages, ambiguity and ultimate conveying different semantics; it is quite complicated to construct the characteristics for bilingual language processing. It must be pointed out that the research on semantic and syntactic types of written language processing is present in linguistic literature, though only focused on English in conjunction with the other languages.

Written bilingual language processing on the sentence level

A previous study (Navracsics & Sáry, 2015) presented the outcome of a psychophysical bilingual experiment carried out on 97 Hungarian speakers of English as L2.

The participants were exposed to L1 and L2 sentences, and they were anticipated to make decisions concerning the acceptability of those sentences. The authors concluded significant differences in L1 and L2 concerning semantic and syntactic types of processing in both RTs and acceptability judgments in most cases. They also found that language proficiency plays a decisive part in the processing of sentences, in contrast to the age of acquisition of L2.

The other study (Ravi & Chengappa, 2014) directed at "exploring the semantic and syntactic processing variances between native and second languages in 20 early high proficient Kannada–English bilingual adults through accuracy and RT measurements." Their subjects participated in a semantic judgment task as well as the task of syntactic judgment both in Kannada and in English. The participants were requested to respond by pressing keyboard buttons. The results showed that the level of accuracy was better in L1 than in L2 for all three types of sentences, whereas the overall processing speed was better in L2 than in L1 for all the given types of sentences. It is indicated that participants handled semantic violations with higher accuracy and speed of processing if comparing with syntactic violations.

The study (Foursha, Austin, & van de Walle, 2006) on grammaticality judgment found that "early, balanced Spanish-English bilinguals produced higher RTs in response to sentences, indicating a processing effect compared to English monolinguals." Nonetheless, on accuracy rates, bilinguals attained like English monolingual speakers. The authors concluded that these results "provided no evidence for cross-linguistic syntactic interference in early, balanced bilinguals during processing of English in comparison with Spanish."

Methodology

Research instruments

The present study used a mixed methods design. To follow the set trend of using a behavioral measurement in an investigation into processing of languages, which has blossomed in the past two decades and with the aim of measuring performance of participants in written language comprehension, the psychophysical method was implemented. The use of a custom-made MATLAB program (MATLAB, 2014) allowed to present the experimental stimuli. The outcome of each trial is the data with reaction times and accuracies, which are the important research strands.

In order to collect data relatively promptly and in a unified fashion, paper-based questionnaire, relating to participants' linguistic configuration, as a research instrument was implemented. Thus, all the participants have filled in information about themselves such as their sex, age, handedness, age at the onset of acquisition of English, way of second language acquisition. The questionnaire asked the participants to give answers to fourteen questions: eight questions of which have three options, five dichotomous questions (to choose between two alternatives) and one open-ended question, requesting information from the participants. Among the advantages of using such an instrument is that it is quite useful for large groups when interviews would be impractical and due to further visual comparison of

answers. The standardized fashion of questionnaire approaches all the participants of the study in the same manner.

Participants

Due to the infeasibility of random selection of the participants, a convenience sampling was employed. Twenty-one Russian-speaking users of English as L2 (19 females and 2 males), who are early bilinguals, have participated in the study. Most of them are undergraduate students of English Studies and Psychology majors at Kalmyk State University named after B. B. Gorodovikov (Elista, Russia), while the rest is researchers at the aforementioned university. The accepted subjects are with self-reported C1 and B2-levels of English in accordance with The Common European Framework of Reference for Languages (CEFR, 2001). As for the students, their midterm and final exam test results implied that there were no outliers with quite high or quite low proficiency levels compared to peers.

The mean age of the participants is 21.8 years old. They were 19 right-handed and 2 left-handed subjects with normal or corrected-to-normal vision. None had a history of reading disabilities or neurological and psychiatric disorders. All the participants gave their written informed consent to participate in the experiment.

Material

The stimulus material is based on the two languages of the Indo-European language family, Russian and English, which are especially dissimilar in various essential facets. In particular, the grammatical structures demonstrate substantial alterations. As such, English has a fairly fixed word order. Meaning is conveyed via the addition of words (for example, auxiliaries) and movement of words in constrained limits. While Russian expresses meaning chiefly with the help of differences in the alignment of words, for instance, by inflections or the addition of prefixes and suffixes. Its word order is very fluid.

The stimulus of the current study consists of 240 English and Russian sentences, among which are 60 Russian and 60 English correct control sentences. Along with these, 30 Russian and 30 English semantically anomalous sentences, which have critical words in initial, internal or final position. In addition, 30 English and 30 Russian syntactically violated sentences, with errors in word order and case agreement, were created.

Having a clear understanding of the amount of visual information processed during a fixation on sentence, such factors as word's frequency, length and ease of integration into the sentence influence, were considered while generating the stimulus material. Accordingly, all the sentences contain from six to ten syllables and the words in the sentences are included in the list of frequently used words. The sentences have a mean length of 4.77 words (SD = 1.35 words). All the English sentences were derived from the bilingual study on Hungarian-English language users (Navracsics & Sáry, 2015) with the authors' consent, whereas all the Russian sentences were generated by the author.

Data collection procedure and analysis

Data collection was divided into three parts. The subjects were run individually in one session lasting approximately 40 minutes. In the first part, the subjects received explanation about the experiment, read and signed their informed consents and received training for the experiment. In the second part, they filled out a questionnaire. In the third part, subjects were seated in a comfortable chair in front of a laptop and introduced to the pilot behavioral experiment. During the pilot trial, all the subjects have completed a practice with 8 sentences that are likely to appear in the actual task. The subjects were asked to push the left or right arrow to judge about the correctness of the sentence, which made them move to the next sentence. The same procedure repeated until each sentence was over. Once they are done with the trial and understand their task, in the actual experiment, the participants were presented with 30 experimental trials.

The viewing distance was set to be the approximately normal viewing distance of a computer screen (~ 50 cm). The trials consist of semantically and syntactically correct and incorrect sentences in both Russian and English. The 240 stimulus sentences were evenly distributed among 8 blocks and randomly intermixed within each block. In other words, the order of the sentences was randomized for each trial and then presented to all participants, though each participant viewed all the sentences. The sentences were appearing in black letters against a white background for the duration of 5000 milliseconds. Proportional font with a relatively small letter size (28 Segoe UI) was used in order to minimize saccadic eye movements during the psychophysical test. The sentences went visually one by one in the center of the computer screen; each one appeared after the centered red asterisk, which serves as a fixation point.

The participants were instructed to read the sentences carefully and to give their responses as fast and as accurate as possible. Their ultimate task was to decide by pushing the right or left arrow of the keyboard whether the sentence is correct or not. If they thought that the sentence is correct, they had to push the right arrow, if the sentence is not correct, they used the left arrow. Beyond time, the computer did not accept decisions, and the fixation point was appearing again for 2000 milliseconds before the next sentence came up. The program recorded correct and incorrect hits alongside response latency times.

Findings and Discussion

The first research question

The reaction time results (see Figures 1 & 2) have shown the following numerical data. In the processing of semantically correct Russian sentences the reaction time is the shortest among all – 1.44 s, followed by syntactically correct Russian sentences – 1.54 s, semantically incorrect Russian sentences – 1.57 s, syntactically incorrect Russian sentences – 1.59 s, semantically incorrect English sentences – 2.03 s, syntactically violated English sentences – 2.04 s, semantically correct English – 2.06 s, syntactically correct English

sentences – 2.09 s. The mean reaction time for semantic comprehension is 1.78 s, while for syntactic it is equal to 1.81 s.

Figure 1. *RTs in comprehension in Russian, s*
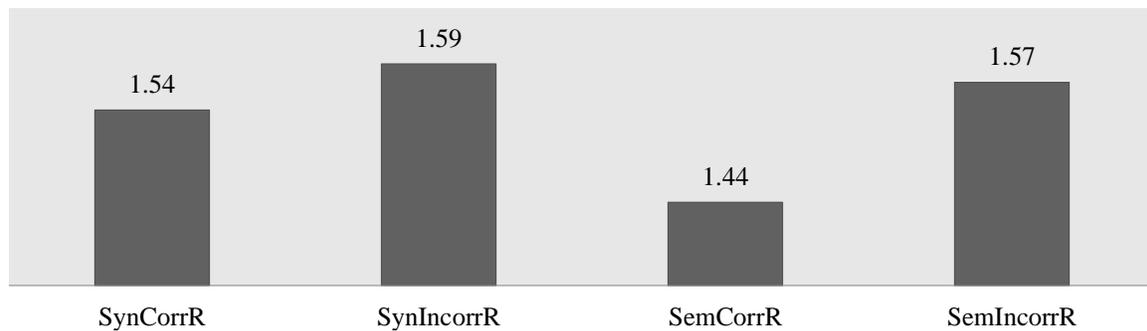
Figure 2. *RTs in comprehension in English, s*
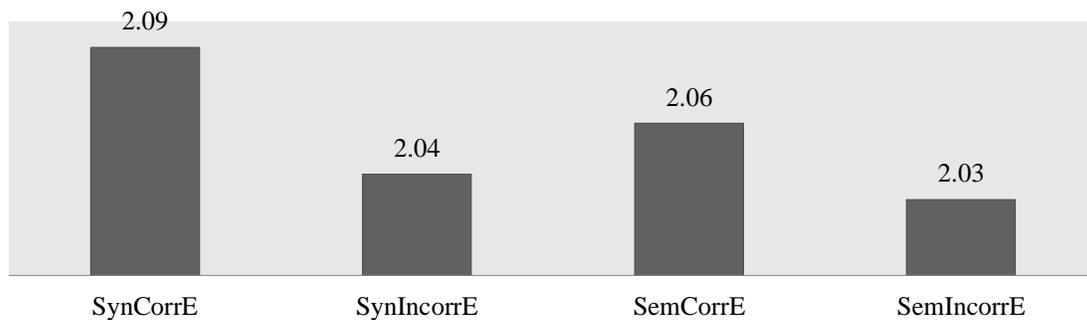
The data with RTs per each type of sentence within the two languages are given in Table 1.

Table 1. *RTs for four types of sentences across L1 and L2*

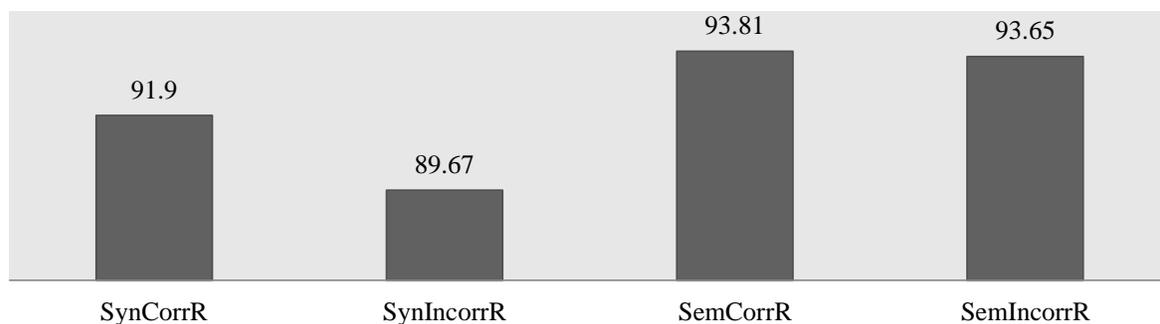| Type of sentences	Language	RT, s
Semantically correct	Russian	1.44
	English	2.06
Syntactically correct	Russian	1.54
	English	2.09
Semantically violated	Russian	1.57
	English	2.03
Syntactically violated	Russian	1.59
	English	2.04

As it can be observed, semantic type of written language processing at the sentence level in L1 and L2 seems to be at around same level as syntactic type in time constraints.

The second research question

Considering the accuracy results, the best were which the participants achieved in the judgments about the Russian sentences in the following order: semantically correct sentences, semantically incorrect sentences, syntactically correct sentences, and syntactically incorrect sentences.

The next results were those of the English sentences: syntactically violated sentences; semantically correct sentences, semantically violated sentences and syntactically correct sentences. For comparing the L1 comprehension across semantics and syntax, see Figure 3. In order to compare semantic and syntactic comprehension in English, see Figure 4.

Figure 3. *Accuracy of judgments in comprehension in Russian, %*

Figure 4. *Accuracy of judgments in comprehension in English, %*
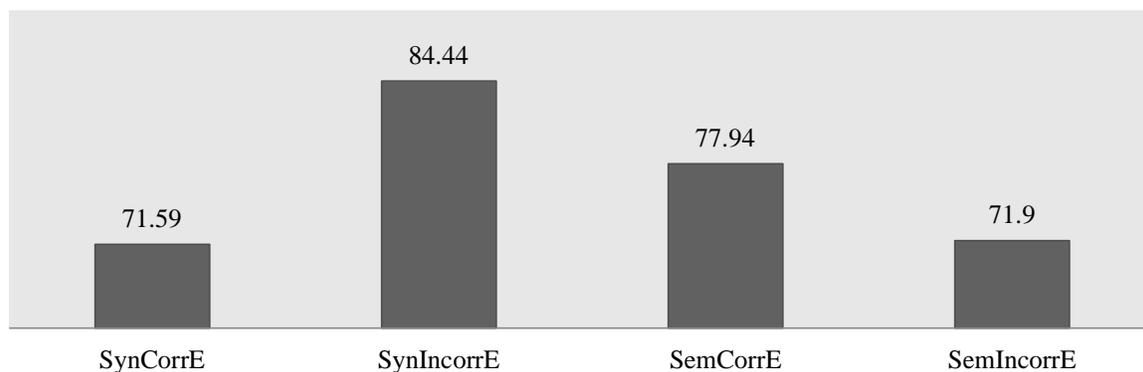
The data with percentages per each type of sentence within the two languages are given in Table 2.

Table 2. *Accuracy for four types of sentences across L1 and L2*

Type of sentences	Language	Accuracy, %
Semantically correct	Russian	93.81
	English	77.94
Syntactically correct	Russian	89.67
	English	71.59
Semantically violated	Russian	93.65
	English	71.90
Syntactically violated	Russian	89.67
	English	84.44

The subjects showed the smallest error rate (6.19%) and shortest response time (M = 2.03 s) in the semantic comprehension, and the highest error rate (28.41%) and longest response time (M = 2.09 s) in the syntactic comprehension. As expected, the subjects generally performed well on the comprehension sentences in their L1, with the mean accuracy equal to 92.25%. Looking at the figures and tables above, it is observed that semantic type of written language processing at the sentence level in L1 and L2 is more accurate than syntactic type.

Questionnaire findings

In accordance with the data obtained from the questionnaire, all the participants speak Russian at home, while 2 participants have specified that they use both languages while communicating with their friends. Mother's L1 is Russian of the participants in almost all the cases, except for one, when father's L1 is indicated as Russian – 18, Kalmyk – 3. Eighteen subjects have mentioned that Russian is the language they are comfortable in, while 2 participants feel comfortable in both (Russian and English), and 1 participant – in English. Fourteen participants read more in Russian, while 7 prefer to do it in English. Books, scientific and Internet articles, comics, magazines are the favorite sources of reading for the participants of the study. Sixteen participants prefer to read books only, 5 participants are fond of reading articles (4 – scientific; 1 – Internet), while two participants are likely to read books along with comics or magazines respectively. Average reading time for the participants is several hours for 9 participants, 1 hour for 7 participants, and 30 minutes for 5 participants. Among the reasons to do so, the subjects have specified enjoyment (19) and its obligatory nature (1), while one of the participants marked both options.

Twelve participants have stated in the questionnaire that they do not consider themselves bilinguals, while the rest of them believe they are. Five of the subjects have spent a certain amount of time in an English-speaking country, 2 months as the longest and 11

days as the shortest. Three of the participants have stayed in an English-speaking country for 3 weeks, while two of them spent 2 weeks and 1 participant stayed for 1 month there. Many of the participants would like to work or study abroad, 18 out of 21 subjects. They are all exposed to English daily, as the younger participants are English Studies and Psychology major university students, and the older subjects are researchers at the university, who more or less use English in their everyday and professional life.

Statistical findings

The received data were primarily checked for normality via SPSS. The test of normality (Shapiro-Wilk) showed that the data is normally distributed (bigger than .05), that is why parametric test was chosen. The within group comparison showed that there is no significant difference between RTs in semantic and syntactic written language processing at the sentence level in L1 and L2 (see Tables 3 & 4).

Table 3. *The results of a paired-samples t-test for RTs in L1*

SynCorrR	SemCorrR	SynIncorrR	SemIncorrR
M= 1.54, SD= .33	M= 1.44, SD= .28	M= 1.59, SD= .30	M= 1.57, SD= .36
t(20)= 3.109, p = .006		t(20)= .575, p = .572	

Table 4. *The results of a paired-samples t-test for RTs in L2*

SynCorrE	SemCorrE	SynIncorrE	SemIncorrE
M= 2.09, SD= .58	M= 2.06, SD= .60	M= 2.04, SD= .57	M= 2.03, SD= .60
t(20)= .874, p = .393		t(20)= .271, p = .789	

It is found that there is a significant difference between accuracies in semantic and syntactic written language processing at the sentence level in L1 and L2 (see Tables 5 & 6).

Table 5. *The results of a paired-samples t-test for accuracies in L1*

SynCorrR	SemCorrR	SynIncorrR	SemIncorrR
M= 91.90, SD= 7.11	M= 93.80, SD= 6.08	M= 88.25, SD= 9.46	M= 93.65, SD= 5.04
t(20)= -1.333, p = .197		t(20)= -2.795, p = .011	

Table 6. *The results of a paired-samples t-test for accuracies in L2*

SynCorrE	SemCorrE	SynIncorrE	SemIncorrE
M= 71.58, SD= 12.76	M= 77.93, SD= 14.27	M= 84.44, SD= 12.26	M= 71.90, SD= 17.53
t(20)= -3.537, p = .002		t(20)= 4.659, p = .000	

The results of the analysis of reaction times address the first research question of the conducted study: it is suggested that semantic type of written language processing at the sentence level in L1 and L2 is not significantly faster than syntactic type. In accordance with the analysis of the data of accuracies, it is argued that semantic type of written language processing at the sentence level in L1 and L2 is more accurate than syntactic type.

Overall, the results of the research validated the BIA model (Dijkstra & van Heuven, 1998) in the sense that "there is a precise mechanism for the way in which orthographic forms are activated in two languages when a bilingual recognizes visually presented words or sentences." As soon as orthographies of Russian and English are not similar, there is a non-parallel activation that results in less competition at the lexical and sub-lexical levels.

Conclusion

Among possible limitations of this study is a rather limited number of employed subjects due to the restricted time constraints of this study. This point might be taken into consideration for future research by the author of this research and possibly by other investigators who are interested in this area. As a suggestion for the further investigation, age and other retrieved information from the questionnaire might be focused more deeply and presumably may lead to varied conclusions.

Disclosure statement

No potential conflict of interest was reported by the author.

Acknowledgments

would like to say big thanks to my Mother, Svetlana Darzhinova, for her immense support and for always believing in me and encouraging me in my journey.

References

Abisheva, K. (2001). Socio-linguistic contactology. *Galym*, 426.

Abu-Rabia, S. (2001). Testing the interdependence hypothesis among native adult bilingual Russian-English students. *Journal of Psycholinguistic Research, 30*(4), 437-455.

Barsalou, L. W. (2008). Grounded cognition. *Annu. Rev. Psychol, 59*, 617–645.

Bråten, I., Britt, M. A., Strømsø, H. I., & Rouet, J.-F. (2011). The role of epistemic beliefs in the comprehension of multiple expository texts: Toward an integrated model. *Educational Psychologist, 46*(1), 48-70.

Council of Europe. (2001). Common European framework of reference for languages: *Learning, teaching, assessment*. Cambridge, U.K: Press Syndicate of the University of Cambridge.

de Groot, A. M. B., Delmaar, P., & Lupker, S. J. (2000). The Processing of Interlexical Homographs in Translation Recognition and Lexical Decision: Support for Non-Selective Access to Bilingual Memory. *The Quarterly Journal of Experimental Psychology Section A, 53*(2), 397–428.

de Koning, B. B., & van der Schoot, M. (2013). Becoming part of the story! Refueling the interest in visualization strategies for reading comprehension. *Educational Psychology Review, 25*, 261-287.

Dijkstra, T., & van Heuven, W. J. B. (1998). The BIA model and bilingual word recognition. In J. Grainger & A. M. Jacobs (Eds.), *Scientific psychology series. Localist connectionist approaches to human cognition* (pp. 189-225). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.

Eberhard, D. M., Gary F. S., & Charles, D. F. (2019). Ethnologue: Languages of the World. Twenty-second edition. Dallas, Texas: SIL International.

Engelen, J., Bouwmeester, S., de Bruin, A., & Zwaan, R. (2011). Perceptual simulation in developing language comprehension. *Journal of experimental child psychology, 110*. 659-75.

Foursha, C. D., Austin, J. B., & van de Walle, G. A. (2006). Is language processing identical in monolinguals and early, balanced bilinguals? In D. Bamman, T. Magnitskaia, & C. Zaller (Eds.), *Proceedings Supplement of the 30th Boston University Conference on Language Development*.

Hagoort, P. (2013). MUC (Memory, Unification, Control) and beyond. *Front. Psychol, 4*, 416.

Hopp, H. (2015). The timing of lexical and syntactic processes in second language sentence comprehension. *Applied Psycholinguistics, 37*(5), 1253-1280.

Jouravlev, O., & Jared D. (2019). Native language processing is influenced by L2-to-L1 translation ambiguity, *Language, Cognition and Neuroscience.*

Kendeou, P., McMaster, K. L., & Christ, T. J. (2016). Reading Comprehension: Core Components and Processes. *Policy Insights from the Behavioral and Brain Sciences, 3*(1), 62–69.

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review, 95*(2), 163-182.

Mallamma, V. R. et al. (2014). Semantical and syntactical Analysis of NLP. *International Journal of Computer Science and Information Technologies, 5*(3), 3236 – 3238.

Marian, V., & Spivey, M. (2003). Bilingual and monolingual processing of competing lexical items. *Applied Psycholinguistics, 24*(2), 173-193.

MATLAB. (2014). Version R2014b. Natick, Massachusetts: The MathWorks Inc.

Navracsics, J., & Sary, G. (2015). Written sentence comprehension in L1 and L2. In: *Multidisciplinary Approaches to Multilingualism*, Peter Lang Verlag, 339-357.

Pulvermüller, F. (2005). Brain Mechanisms Linking Language and Action. *Nature reviews. Neuroscience. 6*. 576-82.

Ravi S. K., & Chengappa S. K. (2015). Sentence processing in high proficient Kannada–English bilinguals: a reaction time study. *International Journal of Multilingualism, 12*, 3, 376-392.

Roberts, L. (2013). Sentence processing in Bilinguals. *Sentence Processing.* 221-246.

Sherkina-Lieber, M. (2004). The cognate facilitation effect in bilingual speech processing; the case of Russian-English bilingualism. *Cahiers linguistics d'Ottawa, 32*, 108-121.

van den Broek, P. (2010). Using texts in science education: Cognitive processes and knowledge representation. *Science, 328*, 453-456.

Wu, C. Y., Vissiennon, K., Friederici, A. D., & Brauer, J. (2016). Preschoolers' brains rely on semantic cues prior to the mastery of syntax during sentence comprehension. *NeuroImage, 126*, 256-266.

Zwaan, R., & Radvansky, G. (1998). Situation Models in Language Comprehension and Memory. *Psychological bulletin, 123*, 162-85.

Zwaan, R., & Pecher D. (2012). Revisiting Mental Simulation in Language Comprehension: Six Replication Attempts. *PLOS ONE 7*(12): e51382.

Biographical notes

LIUBOV DARZHINOVA is currently a researcher and a PhD fellow in Linguistics at the Center for Research on Linguistics and Language Studies, The Education University of Hong Kong. Liubov is also a teacher-volunteer of blended learning course of IELTS at the Center for Language in Education, The Education University of Hong Kong. She completed her MA in Applied Linguistics at the University of Pannonia, Hungary and got her BA in English and German Studies from Kalmyk State University, Russia. Her research interests include bilingualism, language assessment, language processing, second language acquisition, etc.