Review of *Semantic processing in the native and non-native language: The role of working memory and the type of semantic relations*, dissertation presented by Dmytro Khanzhyn at the School of Languages and Literatures of the Adam Mickiewicz University, Poznań, Poland

This is a nice piece of work, which I enjoyed reading a lot. The whole dissertation is written clearly and in an impeccable, easy to read, English.

After a short introduction, the dissertation continues with two introductory chapters, the first one on lexico-semantic processing and semantic priming, the second on working memory. In the next three chapters, six empirical studies are described. These chapters are followed by a thorough general discussion of the findings and a short conclusion chapter. I will describe my evaluation of the different chapters in detail below.

I want to congratulate the candidate on the first two chapters. I am very well acquainted with all literature reviewed and summarized in Chapter 1 (Lexico-semantic processing and semantic priming), a rather broad area in which lots of articles have been published since the early seventies. The candidate has done an excellent job in presenting the essentials of this literature and thus setting the stage well for the empirical work described later in the dissertation.

It was nice to see, for instance, that in section 1.2.1. (Attributional or feature-based models of semantic memory), even the criticism of the so-called theory-based view was attended to. (I never liked the term 'theory-based view on semantic concepts', because one can hardly call it a real model or theory, but as a criticism on feature-based models, this view touched on an essential point: as long as we cannot explain what counts as a feature and how people come up with such features, we miss a crucial aspect.) I also liked section 1.4. (Semantic relatedness task), which nicely reflects on the differences between the often used semantic priming experiments with a lexical decision task and a semantic relatedness task.

I am only roughly familiar with the literature on working memory (Chapter 2). However, again the candidate succeeded very well in summarizing the state of the art and preparing the reader for the experiments described in Chapters 4 and 5.

Chapter 3 (Polish word associations) is right up my alley. I very much appreciate that the candidate took the trouble to gather word associations in Polish. I believe that the work of Simon De Deyne has shown convincingly the important information that can be distilled from such data. It is an elaborate work to gather word associations (starting with finding and motivating participants, but also cleaning up and analyzing the data), but I think it is worth the trouble, and in this dissertation this chapter served the important purpose of validating the material that was used in the experiments of

Chapters 4 and 5. I wonder, however, why the candidate didn't contact Simon, since the Small World of Words project and the available platform could had easily been adapted to incorporate Polish data. I would recommend that the data that formed the basis of Chapter 3 still be integrated in the Small World of Words project, so that they can also be used by other researchers for possible crosslinguistic research. Furthermore, it would yield possibilities for Polish researchers (and interested participants) to enlarge the Polish word association data set. (Needless to say, all credits would be given to the candidate.)

One particular aspect of Simon's and my work with the Small World of Words is not mentioned in Chapter 3. In some of our more recent publications (e.g., De Deyne et al., 2016), we look beyond direct associations between cue and response words. Inspired by the ideas of activation spreading from the early era of cognitive psychology in the seventies, we applied random walk techniques that allowed us to investigate the global structure of the association network. The random walks incorporate activation spreading from, for instance, 'lion' to 'stripes', through the intermediate knot of 'tiger' (assuming the links between lion and tiger and between tiger and stripes are strong enough). In all studies where we applied this technique, we obtained the best results for this 'global structure model'. However, such a random walk process requires a data set that is much larger than the set gathered in Chapter 3 (and, of course, much larger than what one expect from the candidate to gather). Some discussion of this extension of the Small World of Words project would had be justified. If the Polish data set could be enlarged through crowd sourcing by linguists and psychologists in Poland, future work might make use of such a random walk implementation.

It is remarkable that the correlations between corpus-based semantic vectors and R1 similarity, presented in Table 8 at page 86 ($r = 0.35$) is somewhat higher than the correlation for the R123 similarity ($r = 0.24$). I'd be interested to know if the candidate has an idea why, since we invariably find the best predictions for R123, no matter what other variable we predict.

In the dissertation, word2vec is usually used as the standard similarity measure with which association-derived measures are compared. This is in line with what most authors in the field do. I have a different opinion about this, however. Text corpora always consist of information that comes from communication (written or spoken). But communication takes place in a context where sender and receiver have assumptions about what is redundant in the messages they send, because of common knowledge that needs no spelling out. As a result, this information is often a bit biased. Our favorite example is the color of bananas. According to the Small World of Words, yellow is the closest color in the network to banana. In text corpora, however, green and brown are usually closer to banana than yellow, simply because one doesn't mention the color of bananas often if they have the

default value yellow, but when the color is *not* yellow, it is explicitly mentioned. (Apart from this disadvantage, of course, text corpus based measures have the advantage of scale, compared to word associations, where lots of new association data are needed to enlarge the data set.)

Chapter 4 presents a series of three interesting experiments. I will give some thoughts bellow that might be worth some reflection.

First, in the literature on priming, one often sees contrasting conditions that are compared. This is also the case in Chapter 4: pairs are either strongly related, weakly related, or unrelated. I like to think of relatedness as a continuum that could be analyzed as such (using correlational measures or as a random effects variable) instead of treating the three conditions as discrete. If I were a jury member in a public defense, I'd like to hear the candidate's opinion on that.

Second, it might be interesting to get an idea of the reliability of the reaction times (for instance, by splitting up the respondents into two half groups, and then applying the Spearman-Brown reliability calculation.

Third, is it possible that the visual working memory task was just easier than the verbal task and therefore interferes less? I don't think that this simple objection cuts wood, but does the candidate have a clear argument against this counterargument?

Fourth, probably the question I most liked to be answered about Chapter 4 is the following. It is intriguing that weakly related pairs consistently yield the highest response times, significantly higher than the strongly related or the unrelated pairs. My guess is that this is the result of a response process, not so much activation of semantic information perse. Let me spell this out in detail. Think of semantics as a multidimensional space. Semantic distance between word pairs then can be projected unto a dimension that corresponds to 'sematic relatedness'. In the semantic relatedness task, participants have to give a dichotomous response, while the semantic distance is a continuum that varies from highly semantically related (synonyms, for instance) to completely unrelated. So somewhere, the participant has to draw the border between what can still be called 'related' and what is 'unrelated'. In other words, the task of the participant comes down to a categorization task. It is well known in the categorization literature that categorizing stimuli close to the category boundary takes more time than categorizing stimuli further away from the border. (See, for instance, the models of Gregory Ashby and Robert Nosofsky on this topic.) In this view, the longer response times for weakly related word pairs are the result of a process that takes place *after* semantic activation of the two words involved. It would just be the result of 'some hesitation', you could say, of whether less-than-clear-relatedness should be classified as 'related' or as 'not related'. (When does one say

that two words are *not* related?) How does this possible explanation relate to the 'inhibition' process the candidate seems to favor?

Chapter 5 is a nice study with clear results that suggest similar semantic activation mechanisms in both native and non-native speakers. These results should be taken into account by all models about bilinguals. I particularly like the thorough analyses of the data, redoing the analyses without words with higher error rates.

What I liked most about the dissertation is the very thorough General Discussion (Chapter 6), in which all finding are meticulously weighted and related to the literature. I congratulate the candidate on this section particularly. (Very good section 6.3.3. on limitations and future research.)

Concluding, the dissertation presents some nice studies that deserve a place in the literature (besides good summaries on semantic memory and working memory). Also, we can say that the dissertation leaves some questions unsolved and raises new puzzles, as is clearly acknowledged by the candidate in his last chapter. But as often said: new answers point at ever more questions. I want to congratulate the candidate for his work and in my opinion he can surely proceed to the public defense stage. The issues I raised in my evaluation need not be dealt with in detail in his PhD. They are more meant as food for thought on future work (though, of course, I don't mind if he prefers to add comments on some of them in a final version).

Finally, it is difficult for me to judge on the 'cum laude' degree, as I have no comparison. (In our university, degrees for PhD's have been abolished many decades ago.) Given my clearly positive evaluation of the presented work, I would surely not object.

Minor details and typo's:

Page 20, line 5: "Finally, this chapter includes a discussion and uses …" (instead of 'use').

Page 41, line 14: "The primes were presented very briefly …" (delete 'for').

Page 130, beginning of section 5.3.2.2.: Consistent instead of Consistently ?

Page 149, line 18: "Because there is *a* lack of association …" (inset 'a').

Page 152, line 4 from below: "… that follow from *the* Study 1 results …" (insert 'the').

Leuven, August 21 2023

Gert Storms