NetWordS Short Visit Grant Scientific Report

Cynthia Siew

Proposal title: Multiplex language network of phonological and orthographic representations

Application reference number: 6503

*Purpose of the visit*

The purpose of this visit was to gain some insights into building a multiplex network of language to simultaneously represent the phonological and orthographic relationships that exist between words. Constructing a multiplex network with phonological and orthographic layers could lead to a deeper understanding of how the network structure of the phonological network influences visual word processing and how the network structure of the orthographic network influences auditory word processing. The metrics obtained from the multiplex can inspire future psycholinguistic experiments to test hypotheses on auditory and visual lexical processing.

*Description of the work carried out during the visit*

Development and implementation of two network metrics specific to the language multiplex of orthographic and phonological forms: (a) clustering coefficient, (b) average path length/closeness centrality.

There are two types of *clustering coefficient* in this context. The first clustering coefficient refers to the *orthographic* clustering of a target *phonological* word form, *Cop*, and the second clustering coefficient refers to the *phonological* clustering of a target *orthographic* word form, *Cpo*.

*Cop* describes the density of connections occurring among the orthographic representations of the phonological neighborhood of a target phonological word form. *Cpo* describes the density of connections occurring among the phonological representations of the orthographic neighborhood of a target orthographic word form.

Similarly, there are two types of *average path length* in this context. The first average path length refers to the *orthographic* average path length of a target *phonological* word form, *APLop*, and the second average path length refers to the *phonological* average path length of a target *orthographic* word form, *APLpo*.

*APLop* describes the average number of steps/connections between the orthographic representations of the phonological neighbors of the target phonological word form and the orthographic representation of the target phonological word itself. *APLpo* describes the average number of steps/connections between the phonological representations of the orthographic neighbors of the target orthographic word and the phonological representation of the target orthographic word itself.

These metrics were developed as prior work has shown that network structure of complex networks influences how information is spread among nodes. In the language multiplex, these metrics could represent how the network structure of a *different* layer of the multiplex influences the extent to which activation is spread among word representations from a different modality.

*Description of the main results obtained*

(a) Clustering coefficients were obtained for each word in the multiplex by implementing the clustering coefficient metric that was developed during the visit. Both types of clustering coefficients were obtained, *Cop* and *Cpo*—by considering the phonological relationships among orthographic neighbors of the orthographic representation of words, as well as the orthographic relationships among phonological neighbors of the phonological representation of words. The work that remains to be done is to conduct psycholinguistic experiments (such as word naming and lexical decision) to compare the speed and accuracy of processing words of either high or low *Cop*s and *Cpo*s*.*

(b) Some initial results were also obtained for average path lengths, although the metric would need to be further modified to calculate closeness centrality instead, to avoid the problem of “infinite paths” that occur among disconnected components of the graph. Again, psycholinguistic experiments would have to be conducted to compare the speed and accuracy of processing words of either high or low closeness centrality.

*Future collaboration with host institution (if applicable)*

Future collaboration could involve further refining or perhaps building new metrics for the multiplex.

*Projected publications / articles resulting or to result from the grant*

The multiplex metrics developed during this short visit could be used to motivate several psycholinguistic experiments to evaluate the influence of phonological information on visual word processing, and the influence of orthographic information on spoken word processing. The results could lead to further insights into how the network structure of the phonological network influences visual word processing and how the network structure of the orthographic network influences auditory word processing. In particular, as a multiplex allows for activation to spread across representations of words in different modalities, the potential results may reflect a more accurate representation of the dynamics that occur within a language network during lexical processing.

*Other comments (if any)*

During this trip I also completed a short tutorial on the igraph library in R, solved various computational problems encountered while implementing the code to calculate the above metrics in R, and investigated how to make the calculation of the clustering coefficients and average path lengths more efficient.