

# **A Look to the Future: Big Data in Neurorehabilitation**

Yasmeen Farooqi-Shah

The term Big Data is pervasive in industry, cybersecurity, and popular media, and we have heard that it has been transformative in analyzing and predicting behavior. However, the use of Big Data has been less frequent in Speech-Language Pathology. The goal of this issue of Seminars in Speech and Language is to illustrate recent work that uses Big Data to advance the understanding of language breakdown and recovery in adult neurorehabilitation.

The first article in this issue introduces the topic of Big Data and discusses opportunities and challenges in using Big Data in adult neurorehabilitation. The second article in this issue, authored by MacWhinney and Fromm, describes the genesis of the AphasiaBank project and automatic computational analysis tools, which have been used to examine language in aphasia. MacWhinney and Fromm review the findings of 45 studies that used AphasiaBank data to examine a variety of questions in discourse, grammar, gesture, lexicon, fluency, aphasia classification, social factors, and language recovery. The next two articles, one by Thorne and Farooqi-Shah and the other by Malyutina, Richardson, and den Ouden, used AphasiaBank data to examine production of verbs in individuals with aphasia. The significance of these two studies is that prior research on word retrieval in aphasia has been predominantly conducted with nouns, and the few verb-based studies have primarily focused on persons with agrammatic language in small-group studies. Thorne and Farooqi-Shah tested the production of verbs with different semantic complexities (light and heavy verbs) and how this interacts with measures of semantic and syntactic impairment. They found that persons who produce fewer light verbs have poorer syntax, and conversely, a paucity of heavy verbs is associated with semantic limitations. Malyutina and colleagues investigated the syntactic sophistication of verb production by examining production of verb argument structure. They found that variables such as word length and frequency affect verb selection more strongly than argument structure. In the final article, Kiran describes large amounts of data accumulated from language practice tasks implemented via iPad apps. Individual learning profiles gleaned from task-specific improvement and use of hints were examined in the context of aphasia severity and showed that persons more severely affected with aphasia can make substantial improvements although their frequent reliance on hints was not necessarily beneficial for their language outcomes.

The articles in this issue showcase how analysis of Big Data has tested existing theories and provided novel insights into the intricacies of language breakdown and recovery aphasia. It is hoped that this issue will not only highlight the importance and potential of Big Data in advancing theoretical and clinical knowledge, but also stimulate readers to contribute to and benefit from Big Data endeavors in aphasiology. I would like to thank Audrey Holland for inviting us to share our work in this issue of Seminars in Speech and Language.