Title: Introducing Python and R for Language Research and Education

Miran Kim

(Dept. of English Education, Gyeongsang National University)

As we all are aware, the 4th Industrial Revolution brings a fundamental change in the way we learn, teach, and communicate one another as well as in how we access to data and knowledge. To cope with the current digital transformation, it can be also helpful for language researchers and educators to understand what is "coding (with programming language)", why it is so emphasized in educational environments, and what we can do with it for our career. In this direction, this lecture will introduce two open-source programming languages, Python and R, and their applications in language research and education.

According to the R Project webpage (r-project.org), "R is a language and environment for statistical computing and graphics." R has been popular with researchers from many disciplines over 30 years. Language researchers use R codes mostly for natural language processing (e.g., frequency analysis, collocation, lexical diversity, etc.), statistics, and data visualization. Although R and Python are very similar in many aspects, while R is more specialized in statistical analysis, Python (python.org) is created as a general (or multi-) purpose programming language. Python is also well suited for machine learning or deep learning, and its popularity has been growing. In 2022, Python is actually the first ranked (or the second depending on indices) programming language for data science and Artificial Intelligence. For example, *Forbes* wrote an article, already in 2014, "*The best 8 languages to choose for learning programming in a month*", and it said, "the language (Python) is easy to learn, handy to read and write, and extremely flexible." This prediction became real and probably this is why most students start learning coding with Python nowadays. On the other hand, the article picked R for "useful sophisticated analytical and statistical tools." R and Python are very similar in many aspects and there are strengths and weaknesses. However, with the recent online implementation of coding platforms (e.g., Colab, RStudio, etc.), we are able to make the best use of both programming languages for our purposes, such as handling big data including unstructured data and graphic data, or employing teaching strategies for more personalized, creative, advanced (including AR, VR, AI) and dynamic learning activities.

By sharing my learning experience of Python (and R) and some of my coding applications implemented to assist English language learners, it is hoped that this lecture can get audience's attention to these amazing tools for language research and education.