

Ανάλυση Λαθών στον Τυποποιημένο Λόγο

Υιοθετώντας ως μεθοδολογία τις αρχές της υπολογιστικής ανάλυσης λαθών (Dagneaux et al. 1998, Granger, 2008, Lüdeling & Hirschmann 2015) η παρούσα ανακοίνωση παρουσιάζει την αποτύπωση των λαθών που διέπουν τη χρήση στερεοτυπικού λόγου, όπως αυτός παρουσιάζεται στις γραπτές παραγωγές μαθητών και μαθητριών της ελληνικής ως δεύτερης/ξένης γλώσσας στο Ελληνικό Σώμα Κειμένων Μαθητών/Μαθητριών ΙΙ (ΕΣΚΕΙΜΑΘΙΙ).

Το ΕΣΚΕΙΜΑΘΙΙ είναι το μεγαλύτερο Σώμα Κειμένων Μαθητών και Μαθητριών για τα ελληνικά ως ξένη/δεύτερη γλώσσα και αποτελείται από περίπου 300.000 λέξεις, ενώ εμπλουτίζεται συνεχώς με νέες γραπτές και προφορικές παραγωγές. Παράλληλα συλλέγονται και λεπτομερή στοιχεία (μεταδεδομένα) που συγκροτούν το γλωσσικό προφίλ των συμμετεχόντων/-ουσών, όπως χώρα προέλευσης, επίπεδο γλωσσομάθειας, μητρική γλώσσα κ.λπ.

Μέσω του συνδυασμού της Υπολογιστικής Ανάλυσης Λαθών και της Γλωσσολογίας Σωμάτων Κειμένων θα γίνει μια προσπάθεια απάντησης σε ερωτήματα, όπως πώς συνδέεται η χρήση του τυποποιημένου λόγου με τις επιδόσεις μαθητών και μαθητριών; Σχετίζεται η λανθασμένη χρήση του με παραμέτρους, όπως μητρική γλώσσα, ηλικία, επίπεδο γλωσσομάθειας, χρόνο εκμάθησης, τρόπο έκθεσης στη γλώσσα; Ποια προβλήματα προκύπτουν κατά την επισημείωση λαθών στον τυποποιημένο λόγο και πώς επιλύονται;

Σε μια προσπάθεια επίλυσης των παραπάνω ακολουθείται η προσέγγιση των Αναστασιάδη & Ευθυμίου (2006) & Μαρκαντωνάτου (2017) στην οποία η στερεοτυπικότητα οριοθετήθηκε βάσει γλωσσολογικών κριτηρίων και εν συνεχεία δημιουργήθηκε ένα λεπτομερές πλαίσιο επισημείωσης λαθών του στερεοτυπικού λόγου των μαθητών και μαθητριών είτε πρόκειται για παραγωγή σωστών δομών είτε για λανθασμένων. Το συγκεκριμένο πλαίσιο επισημείωσης περιλαμβάνει πληροφορίες τόσο την ταξινόμηση του ίδιου του φαινομένου (Τυποποιημένος Λόγος) όσο και την ταξινόμηση των λαθών βάσει σημασιολογίας, λεξιλογίου, μορφοσύνταξης, φωνητικής και πραγματολογίας.

Αρχικά, αφού παρουσιαστούν τα βασικά χαρακτηριστικά του ΕΣΚΕΙΜΑΘΙΙ, στην παρούσα ανακοίνωση επιλέγουμε μια διαφορετική προσέγγιση για τη χαρτογράφηση του στερεοτυπικού λόγου των μαθητών και μαθητριών της ελληνικής ως ξένης/δεύτερης γλώσσας, όπως αυτή περιγράφηκε από τη μέχρι τώρα βιβλιογραφία. Συγκεκριμένα, χρησιμοποιώντας τη στατιστική γλώσσα προγραμματισμού R για τον υπολογισμό του στατιστικού δείκτη λεξικής βαρύτητας (lexical gravity), όπως αυτός αρχικά παρουσιάστηκε από τους Daudaravičius & Marcinkevičienė (2004) και εξελίχθηκε από τους Gries & Mukherjee (2010), θα παρουσιάσουμε μια σειρά μοτίβων της διαγλώσσας (Granger 2019, Raquot & Granger 2012, Wray 2002) τα οποία προέκυψαν από τις γραπτές παραγωγές του ΕΣΚΕΙΜΑΘΙΙ. Ο δείκτης λεξικής βαρύτητας επιτρέπει στον ερευνητή να ανιχνεύσει τις ιδιοματικές φράσεις (collocations) όχι μόνο βάσει συχνότητας εμφάνισης, όπως συμβαίνει με άλλα εργαλεία στατιστικής ανάλυσης (Daudaravičius & Marcinkevičienė, 2004), αλλά και βάσει ποικιλίας του γλωσσικού περιβάλλοντος εμφάνισης της εκάστοτε λέξης/δομής. Στη συνέχεια τα μοτίβα αυτά θα ταξινομηθούν και θα αναλυθούν βάσει ποιοτικών γλωσσολογικών κριτηρίων.

Ως αποτέλεσμα προκύπτει η περιγραφή μιας ολοκληρωμένης διαδικασίας της Υπολογιστικής Ανάλυσης Λαθών σε ένα φαινόμενο και η επεξεργασία της διαγλώσσας των μαθητών και των μαθητριών της ελληνικής ως ξένης ή δεύτερης γλώσσας.

ΒΙΒΛΙΟΓΡΑΦΙΑ

- Dagneaux, E., Denness, S., & Granger, S. (1998). Computer-aided error analysis. *System*, 26 (2), 163–174.
- Daudaravičius, V. & Marcinkevičienė, R. (2004). Gravity counts for the boundaries of collocations. *International Journal of Corpus Linguistics*, 9 (2), 321–348.
- Granger S. (2019). Formulaic Sequences in Learner Corpora. Στο Anna Siyanova-Chanturia and Ana Pellicer-Sánchez (επιμ.) *Understanding Formulaic Language. A Second Language Acquisition Perspective*. New York: Routledge publications. 228-241.
- Granger, S. (2008). Learner corpora. Στο A. Lüdeling & M. Kytö (επιμ.), *Corpus linguistics. An international handbook*, 1. Berlin and New York: Mouton de Gruyter. 259-275.
- Gries, S. & Mukherjee, J. 2010. Lexical Gravity across Varieties of English: AnICE-Based Study of n-Grams in Asian Englishes. *International Journal of Corpus Linguistics* 15. 520–548.
- Lüdeling, A., & Hirschmann, H. (2015). Error annotation systems. Στο S. Granger, F. Meunier, & G. Gilquin (επιμ.), *The Cambridge handbook of learner corpus research*. Cambridge: Cambridge University Press. 135-137.
- Paquot, M. & Granger, S. (2012). Formulaic language in learner corpora. *Annual Review of Applied Linguistics*, 32, 130-149.
- Wray, A. (2002). *Formulaic Language and the Lexicon*. Cambridge: Cambridge University Press.
- Αναστασιάδη-Συμεωνίδη, Α. & Ευθυμίου, Α. (2006). *Οι στερεότυπες εκφράσεις και η διδακτική της νέας ελληνικής ως δεύτερης γλώσσας*. Πατάκης: Αθήνα.
- Τάντος, Α., Μαρκαντωνάτου, Σ., Αναστασιάδη-Συμεωνίδη, Α. & Κυριακοπούλου, Π. (2017). (2η έκδοση). *Υπολογιστική Γλωσσολογία: Θεωρία και πράξη με ειδική αναφορά στην ελληνική γλώσσα*. Ελληνικά Ακαδημαϊκά Ηλεκτρονικά Συγγράμματα και Βοηθήματα.