

## Assessing the Vocabulary of Arabic Bilingual Speakers in Israel

Eihab Abu-Rabiah<sup>1</sup>

<sup>1</sup>Ben-Gurion University of the Negev, Hebrew Language Department  
aehab@post.bgu.ac.il

**Introduction:** In this talk I will present my research on the Hebrew vocabulary of Arabic-speaking bilinguals in Israel. Arabic speakers constitute about 21% of the total population in Israel. They are constantly exposed to Hebrew, the official and majority language in the country, prevailing in government offices, medical institutions, academic institutions and generally in the public space. As such, it is imperative for daily life (Amara, 2017). Currently, Israel has two parallel education systems: one conducted in Hebrew; and the other conducted in Arabic, with Hebrew taught as a required L2 starting from second grade. **Methodology:** This study examines the lexical density (LD) for the first time in expository essays of 23 intermediate Arabic-speaking public high school students over one year (tested in 11th grade and 12th grade) and aims to adapt LD to better suit the morphosyntactic and orthographic properties of Hebrew. LD is one of the most common measures in evaluating progress in vocabulary acquisition in both L1 and L2 (Lauren, 2002). It measures the percentage of content words out of all words (content and function words) in spoken or written texts (Ure, 1971; Berman, 2008). This measure is an important indicator of textual richness because content words convey the bulk of the semantic content of the text (Berman et al., 2011). Testing LD in Hebrew, however, is somewhat challenging because of the orthographic characteristics of this synthetic language: one orthographic word in Hebrew may contain a content word and one or more function words. For example, the single orthographic word שֶׁבֵּיתוֹ (še-be-beto) ‘which is in his house’ parallels five orthographic words in English. To test LD in Hebrew I proposed an adapted approach which enables compatibility with the parallel measure, as used conventionally in analytic languages like English. I used both approaches in testing LD in Hebrew of the Arabic-speaking students, then compared the two results. The main two advantages of the adapted approach are an adequate representation of individual words in the text, in terms of both functions and of frequency. **Findings:** The results show a statistically significant difference between the conventional approach and the adapted approach. No significant increase in LD over the tested period was found, but there was a small effect size. In addition, it was found that with age the students used more morphosyntactically complex parts of speech.

This work was supported by The Ministry of Science and Technology of Israel [scholarship no. 3-16617]

### References

- Amara, M. (2017). *Arabic in Israel: Language, identity and conflict*. Routledge.
- Berman, R. A. (2008). The psycholinguistics of developing text construction. *Journal of Child Language*, 35(4), 735-771.
- Berman, R., Nayditz, R., & Ravid, D. (2011). Linguistic diagnostics of written texts in two school-age populations. *Written Language & Literacy*, 14(2), 161-187.
- Laurén, U. (2002). Some lexical features of immersion pupils' oral and written narration. *Working Papers in Linguistics*, 50, 63-78.
- Ure, J. (1971). Lexical Density and Register Differentiation. In G.E. Perren, & J.M. Trim (Eds.), *Applications of Linguistics* (pp. 443-452). Cambridge University Press, London.