

**The Impact of Diglossia on Acquisition of Spoken and Standard Arabic:
Narrative and Expository Text Production in Speaking and Writing**

Suha Sabbah Shibel

School of Education

Ph.D. Thesis

Submitted to the Senate of Bar-Ilan University

Ramat Gan, Israel

November, 2020

Abstract

Speakers of Arabic as mother tongue grow into a linguistic situation of diglossia, that is, a linguistic duality comprised of two varieties of the language: Spoken Arabic and Standard Arabic. Spoken Palestinian (PA) is Israeli Arab children's mother tongue used for every day discourse, while Standard Arabic is the written language (MSA-W), acquired in the first grade as part of their formal education (Saiegh-Haddad, 2012). There is a linguistic distance between the two language varieties through the whole range of the language components: phonological, lexical, morphological and syntactic (Saiegh-Haddad, 2017; Saiegh-Haddad & Henkin, 2014). This duality challenges Arab speakers throughout the development of their phonological, linguistic and literacy awareness and their ability to read in Arabic, and affects language acquisition and academic achievements later in life. Therefore, this issue raises much research interest in recent decades.

The current study examines the effect of diglossia from the perspective of literate fluent discourse of Arabic speakers, as expressed in language richness in a produced text. In addition, it investigates the differences and the connection between language abilities as expressed in language richness in PA, oral Standard Arabic (MSA-S) and MSA-W.

The study is underpinned theoretically and methodologically by studies from various fields of research. The first is an international cross-linguistic study focusing on the development of oral and written discourse (Berman, 2004; Berman & Ravid, 1999; Berman & Verhoeven, 2002) that served as a model and supplied research tools for the current study. The other field of research focuses on the Arabic language with its unique diglossia in the areas of narrative abilities and development, language literacy and the contribution of speech skills to the learning of writing, and its development in different genres among Arab speaking children.

The study has three main purposes. The first is to examine the differences between linguistic abilities (at micro level) in text production by variety (PA and MSA), by modality (oral text vs. written text), by variety/modality (PA, MSA-S and MSA-W), and by genre (narrative text vs. expository text). The second purpose is to examine how the diglossia and the linguistic differences between PA and MSA show in text production both lexically and in the aspect of fitting the vocabulary to different communication objectives. The third purpose is to check whether the ability to produce a text in PA and in MSA-S predicts text production in MSA-W, which is the ultimate goal of literacy acquisition.

The research predictions concerning the first research question relate to aspects of diglossia in Arabic: variety, modality, variety/modality and the aspect of genre. The first prediction (variety hypothesis) is that higher levels of linguistic richness will emerge in the spoken variety, acquired first as mother tongue. The second prediction (modality hypothesis) is that higher linguistic richness indicators will appear in the MSA-W modality, as compared to the MSA-S modality, owing to the constraints involved in the production of an oral text. The third prediction (the variety/modality hypothesis) is that there will be differences in linguistic richness by both modality and variety, and that MSA-S will present characteristics of both PA and MSA-W, and have an intermediate status between PA and MSA-W. The fourth prediction (genre hypothesis) is that narrative texts will show higher linguistic richness indicators, and that 4th grade students will demonstrate good control of this usually early-acquired genre.

Two predictions were examined regarding the second research question. The first is that PA text production will show use of spoken lexical characteristics, that is, words unique to PA and cognate words. Moreover, such use will also appear in MSA in both modalities, but the MSA-S condition will contain more such words than the MSA-W condition. According to the second prediction, owing to the academic and formal status of the expository genre, there will be more use of MSA lexical characteristics, that is, cognate words and words unique to MSA in expository texts as compared to narrative

ones, and more so in MSA-W as compared to MSA-S. Both predictions emanate from the premise that young children do not have yet full control over their linguistic repertoire and cannot yet adjust the language to the communicative situation in a satisfactory manner.

The prediction concerning the third question is that a connection will be found between the levels of linguistic richness in the narrative and expository genres in PA and MSA-S, on the one hand, and levels of linguistic richness in MSA-W. It is expected that MSA-W linguistic abilities will be explained by PA and MSA-S abilities. This prediction is based on the connection between oral linguistic abilities to literacy development.

The study participants comprised 87 4th grade students aged 9-10 (42 boys and 45 girls), from three state-elementary schools in Kafar Kara (3 classes in each school), all using Palestinian Arabic as mother tongue in the dialect of Israel's center region (Northern Triangle dialect). The participants are all of medium and medium-high socio-economic status, of standard learning abilities.

The study assignments were to produce texts according to the research aim, as following: each participant met the researcher twice. At the beginning of each of the two sessions, he or she watched a short silent film showing conflicts between students, to inspire the production of three texts. In the first session, the student produced three narrative texts (in PA, MSA-S and MSA-W). In the second session, the student produced three expository texts (in PA, MSA-S and MSA-W). There were 522 texts collected, 261 of which were narrative texts and 261 were expository texts, 174 in PA and 348 in MSA – of equal numbers in MSA-S and in MSA-W. In order to check for differences in micro-level linguistic abilities of text production, we chose three indices customary in the research field for measuring linguistic richness levels in Arabic. The indices are (1) linguistic productivity, (2) lexical density and (3) lexical diversity. Each of these indices includes sub-indices according to word-form categories.

The research findings present a new body of knowledge, unstudied so far, about the way the unique Arabic diglossia is expressed in text production by young children in the two language varieties, written and oral. The findings come out of the diverse linguistic-richness indices: the linguistic productivity, density and diversity indices, and the effect of each one of them on the research variables: language variety, modality, variety/modality and genre. We emphasize that, owing to the Arabic diglossia, the effect of the variety can be distinguished from that of modality (unlike in most other languages), since MSA can be expressed in two modalities, written and spoken, while PA has a spoken modality only. Therefore, we could indicate differences in linguistic richness under the effect of the variety of PA vs. MSA, as well as compare MSA-S vs. MSA-W under the effect of modality.

The findings show that linguistic richness is affected significantly by variety. This finding was apparent in the following indices: productivity index of clause average length, the general density index, index of density of lexical verbs out of all verbs in the text, noun diversity index, and noun morphological structure diversity index. In accordance with our prediction, the general density index was higher in PA, that the children control already at this young age, which enables them to provide more content, and not only language information, owing to easier availability of their cognitive abilities (memory, attention etc.). However, all the other indices were higher in MSA as compared to PA, contrary to our expectations. That is to say, although producing an MSA text is harder for children of this age, they produced more complex texts in MSA, although it is less familiar to them, and not mastered yet. The findings indicate that the children perceive MSA as an official higher-level language. These findings support the ‘developmental paradox’ hypothesis (Berman & Nir-Sagiv, 2007) and suggest that children adjust language to the communicative context already at an early age.

The findings also show that linguistic richness is affected by modality. The following indices: the lexical diversity, that is, the general diversity of the text index, verb diversity and lexical verb diversity showed higher levels in MSA-W as compared

to MSA-S, in accord with our expectation. These results support the effect of modality alone, and not of the variety (since the variety here is identical), or of what in other languages is connected to register.

In addition, the study indicates as well that linguistic richness was affected by variety and modality simultaneously, and that there were differences between the three conditions: PA, MSA-S and MSA-W. This was demonstrated by the following indices: number of clauses, number of types, verb density, density of nouns out of the sum of content words, verb form diversity and verb root diversity. These findings indicate the intermediate status of MSA-S. As to the effect of variety/modality, it was found to affect a higher level of linguistic richness in relation to the levels of linguistic richness affected by either modality or variety separately. The findings pointing to the significant differences in linguistic richness in the context of the triadic comparison of PA, MSA-S and MSA-W highlight the status of the MSA-S as having an intermediate status between PA and MSA-W.

Furthermore, the study shows the effect of genre on textual linguistic richness. This is expressed by the following indices: clause-number productivity index, verb density, number of lexical verbs and adjectives out of all content-words indices, general diversity index and grammatical verbs and adjectives index. The results show more richness in the narrative genre as compared to the expository genre. This finding supports the stance that the narrative structure forms at a very early age. Thus, 4th grade children already produce good standard personal stories. The linguistic richness found in the various verb indices is related to the structure of narrative discourse characterized by recounting occurrences, and by the function of verbs as depicting and expanding events and actions.

However, a large number of linguistic richness indices were higher in the expository genre, in comparison to the narrative one. These indices are as following: clause average length index; index of lexical verb density out of all the verbs in the text; of noun density out of all the words in the text and adjective density out of all content words; noun and adjective density out of all content words indices, and indices of

diversity of verbs, verb forms, verb roots and adjectives. These findings indicate that 9-10 year olds recognize the need to recruit higher language capacities in the expository genre, perceived as formal and related mostly to written school assignments, although, generally, they have not yet reached mastery of this genre.

Regarding the second research question, concerned with the differences in the level at which diglossia is expressed in the production of a text, findings show that children use the language expected from them in a given communicative situation. They use PA (words unique to PA and cognates) when use of PA is required, and use MSA (words unique to MSA and cognates) when use of MSA is required. These findings indicate that children at this age can fit the language to the communicative situation. Yet, the findings show wide use of spoken lexical characteristics (words unique to PA and cognates) in text production in MSA, with wider use of these characteristics in MSA-S than in MSA-W. In addition, there was wide use of MSA lexical characteristics (words unique to MSA and cognates) in text production in PA, though this use was lower than in MSA. Thus, the findings suggest that the children are still in the process of acquiring the diglossia, and are not yet skilled in distinguishing between the varieties and controlling them according to the communicative situation. Therefore, they mix the codes in different situations. The reasons for that might be related to lexical language capacities or cognitive and metacognitive abilities.

As to the third research question, concerning the connection between the indices of linguistic richness in the narrative and expository genres in PA and MSA-S on the one hand, and linguistic richness in MSA-W on the other hand, a positive significant connection was found. In other words, the higher the level of linguistic richness in PA and MSA-S, the higher the linguistic richness of MSA-W.

The findings of the current study offer an important research contribution, both theoretical and practical. Methodologically, this study presents a triadic comparison, between varieties (PA and MSA), and a comparison of modalities (MSA-S and MSA-W). These comparisons have been found to be efficient for understanding the developing

linguistic abilities of children. In addition, the study expands extant knowledge in the field of diglossia research and the effects of diglossia on language acquisition, and sheds light on the linguistic differences between PA and MSA in the interaction of modality and genre.

An additional theoretic contribution of the study is the validation of the linguistic density and diversity indices along those of language productivity, as indices that enable the investigation of linguistic differences by modality, variety and variety/modality.

As to the research practical contribution, the study results have important ramifications on the teaching and assessing of the Arabic language among speakers, owing to the differences in linguistic abilities that the study points to, not only between PA and MSA, but between MSA-S and MSA-W as well. Furthermore, the study is important in that it points to the language abilities in the two spoken varieties for establishing the linguistic phonological, morphological and lexical representations in MSA-W, which is the highest literacy objective of Arabic language teaching.