

# Assessing Arabic

Atta Gebril

*American University in Cairo, Egypt*

Hanada Taha-Thomure

*Bahrain Teachers College, Bahrain*

## Arabic Language: Current Status and Policy

Arabic is spoken as a native language by over 300 million people on two different continents: Africa and Asia. In addition, Arabic is the official language in 26 countries around the world. Given the enormous number of native speakers and the large number of countries where it is spoken as a native language, Arabic was selected as one of the six official languages in the United Nations. Another point of pride for the Arabic language is its historical connection to Islam, since it is the language of the Qur'an. This makes it the religious language of over a billion Muslims around the world. In Muslim countries, learners of Arabic are revered and enjoy several privileges because of this prestigious status of Arabic in Islam (Suleiman, 2003). Recently, there has been a substantial increase in the number of Arabic learners in Western countries. For example, in the USA, numbers of Arabic learners increased from 5,505 in 1998 to 35,083 students in 2009, according to a survey conducted by the Modern Language Association (Furman, Goldberg, & Lusin, 2010). Because of this growing interest, Arabic has become the eighth most studied language in US universities. The renewed interest could be interpreted within a wider context of Westerners attempting to better understand Islam and Muslims. Also, it is clear that there is a growing political and economic interest in the Middle East. More importantly, many Western governments now have a firm belief that learning strategic languages, such as Arabic, is part of their national security strategy, particularly after the events of 9/11 (National Research Council, 2007).

After reclaiming their independence in the second half of the 20th century, the Arab countries espoused educational policies that promoted the teaching and learning of Arabic in schools and the use of the language in the wider society. Arabic was also stated in all Arab constitutions as the national and official

language of these countries. One of the main goals of any educational system in the Arab world has apparently become the training of citizens who can use Standard Arabic accurately and fluently. The purpose of this policy was to revitalize the status of Arabic which was negatively affected by colonial policies that promoted native languages of the imperial powers (see Daoud, 1991, for a comprehensive discussion of the Arabization policies). Modern Standard Arabic, which is a simplified version of Arabic, replaced Classical Arabic in schools and media. Also, Arab governments have provided substantial funding for developing new Arabic curricula and also establishing teacher-training programs. Arabic is currently taught in schools as a core subject from kindergarten to secondary level for around six to eight hours per week. In addition, university students are required to take Arabic as part of their academic programs.

## **Description of the Arabic Language**

### **Linguistic Features of Arabic**

Arabic is a Semitic language that consists of 28 letters written from right to left. Modern Standard Arabic (MSA) is the media language and what is used in the formal public domain. MSA is understood by anyone who has a basic level of education across the Arab world (Taha-Thomure, 2008). It coexists with tens of national and local dialects or vernaculars such as Egyptian, Lebanese, Syrian, Yemeni, Tunisian, and many others that were distributed and redistributed mainly due to large-scale migrations (Owens, 2003). This could be seen as a continuum: MSA resides on one side, and the various vernaculars are dotted all over it. Following Blanc's (1960) approach, one could see the plausibility of this continuum, which includes five levels of speech ranging from plain colloquial to standard classical. Each of those levels has its own linguistic characteristics and features.

### **Diglossia in Arabic**

Charles Ferguson was the first to coin the term "diglossia" in 1959. Ferguson defined diglossia as a language environment in which, in addition to several primary dialects, there is a highly codified variety reserved for written literature and learned largely through formal education. That is the variety of choice for most written and formal spoken purposes; however, it is not usually used by the community for everyday conversation (Ferguson, 1959).

Every native speaker of Arabic acquires his or her own regional/local dialect, which is considered less prestigious (L) within this diglossic context. MSA, or the more prestigious high (H) variety of the language, is taught in schools to all students as of grade 1 across the Arab world. MSA is considered to be a "superposed" variety, in Ferguson's words, and is considered to be the only true form of Arabic (Ferguson, 1959). All written communication is done in MSA, including newspapers, magazines, official letters, formal speeches, bills (both utility bills and government resolutions), and so forth, although recently and with the advent of various Internet tools such as Facebook, Yahoo! groups, Google Chat, and Skype, many have started using dialects as a writing genre at this informal level.

## **Teaching and Learning of Arabic in Schools, Colleges, and the Workplace**

### **Teaching Arabic as a First Language in the Middle East: Realities and Problems**

In most public schools in the Arab world, and in many of its private schools, Arabic language skills are taught through textbooks provided by the ministries of education (Taha-Thomure, 2011). Arabic language textbooks are made up of texts written or edited by Ministry of Education curriculum specialists. These texts often range from those composed of a few words and sentences, to relatively higher-level texts composed of paragraphs, to extended texts that may be one to three pages long at most. Examples of topics include family, school, environment, animals, folk stories, poetry, and myths, which are often edited by authors as needed.

The learning and teaching of Arabic language in most schools have been closely synonymous with textbooks. Many unknowingly believe that the textbook is the school “curriculum” while, in fact, any curriculum must include: standards for teaching, benchmarks and performance indicators, teaching techniques, and assessment tools. The absence of those and of print-rich classrooms has led schools to having a teacher- or textbook-centered curriculum that tests rote learning and knowledge level skills rather than actual achievement and growth (Taha-Thomure, 2008).

### **Teaching Arabic as a Foreign Language in Different Parts of the World: Realities and Problems**

Foreign languages such as Arabic have experienced unprecedented growth in student enrollment—126.5% between 2002 and 2006 (Furman et al., 2010). Such growth has led to the mushrooming of hundreds of Arabic language programs around the USA, where the language is taught primarily by native speakers who do not necessarily have the training needed in pedagogy to effectively teach and assess students’ language proficiency. Many experts in the field of Arabic language teaching complain of the scarcity of Arabic language teachers who are able to combine a solid foundation in pedagogy, teaching methodologies, reflective techniques, and instructional technology (McCarus, 1992). Many initiatives have been launched to help ease those issues including the StarTalk program, which trains thousands of teachers of critical languages including Arabic every summer across the USA (StarTalk, 2011).

## **Assessment of L1/L2 Arabic**

### **Assessment of Arabic as a Foreign Language (AFL)**

This section addresses the AFL assessment practices with specific focus on the US context. One of the most commonly used assessment frameworks in a number of

AFL contexts is the Interagency Language Roundtable (ILR) scale. For example, the ILR Oral Proficiency Interview (OPI) is used by different governmental agencies to make hiring and promotion decisions (Swender, 2003). Also, it is used in intensive AFL programs to make various decisions including ones regarding placement, exit, and admission. A widely used test that was developed based on the ILR platform is the Defense Language Proficiency Test 5 (DLPT 5). The DLPT 5 is the latest version of this test originally developed in the 1950s (Defense Language Institute [DLI], 2012). The test measures three areas: listening comprehension, reading comprehension, and speaking proficiency (using an OPI protocol). It uses a multiple choice format for both the reading and the listening sections, and an interview procedure for the speaking part. The DLPT 5 was mainly developed as a computer-based test (CBT), but it is also available in a paper and pencil format, and the scores are reported according to the ILR scale levels (e.g., 0+, 1, 1+, 2, etc.). Scores obtained from this test are used to assess the language proficiency of military personnel, and decisions regarding promotion or extra pay are made accordingly (DLI, 2012).

In academic circles, most Arabic language tests are developed based on the ACTFL guidelines, which were developed in the 1980s by the American Council on Teaching Foreign Languages (ACTFL) and the Educational Testing Service (ETS). One of the most commonly used exams in this context is the ACTFL Oral Proficiency Interview (OPI), which is an adaptive test between an ACTFL-certified interviewer and an interviewee lasting between 10 and 30 minutes (Swender, 2003). The interviewer asks a number of questions that target different language functions in order to establish the ceiling and floor of the examinee's proficiency. Once she or he decides on a level (out of 11 ACTFL levels), the interview comes to an end. The speech sample is second-rated by another examiner and in case of disagreement a third rater is employed (see Swender, 2003, for more information). The ACTFL OPI is used for a number of purposes, such as placement, diagnosis, and certification of teachers (Gebriel, 2009a). A new version of ACTFL OPI that was recently developed is called ACTFL<sup>®</sup> Oral Proficiency Interview-computer (OPIC). This computer-based interview is conducted between the examinee and a virtual avatar. The responses are digitally recorded and simultaneously stored on a secure electronic system.

The Center for Applied Linguistics (CAL) has a suite of Arabic language tests that are used for a number of purposes in different contexts (Center for Applied Linguistics, 2007):

- *Arabic Proficiency Test (APT)*: This is a paper and pencil test that measures both reading and listening based on the ACTFL levels. APT is used in different contexts including high school, university, and nondegree programs.
- *Arabic Speaking Test (AST)*: This test includes a simulated OPI that measures speaking proficiency in Arabic. AST is tape-mediated since students listen to prompts from a tape and then their responses are simultaneously audiorecorded.
- *Online Arabic Proficiency Test (O-APT)*: O-APT is a paper and pencil test that includes three sections: listening, reading, and writing. The O-APT is used by students who are interested in college credit for the Arabic language.

Mahdi Alish, while working at Ohio State University, developed Arabic Course Achievement Tests that are used as quizzes, midterms, and finals at different levels including beginners, intermediate, and advanced. Also, he developed the Arabic Reading Proficiency Test, which is used with high school, college, and nondegree programs to assess reading proficiency based on ACTFL levels (Center for Applied Linguistics, 2007).

### Assessment of Arabic as a First Language

The educational system in Arab countries is very centralized: Schools use the same textbook and follow the same syllabus. This centralized system has affected assessment practices in schools. Final exams are developed by educational directorates in each governorate and administered on the same day for all students in this region. The centralized system reaches its peak in the high school leaving exam which is a nationwide test. This exam is used in most Arab countries as both an exit and a university admission test, which makes it a very high stakes test. Hargreaves (1997) accurately refers to the stress and pressure students, parents, and teachers go through when preparing for the high school leaving exam in Egypt, and by default in many other Arab countries. Almost 15 years after Hargreaves wrote her description, the situation in Egypt and in most Arab countries is still the same. Assessment in this region has a huge impact on curricula and other school activities. In a study conducted in Jordan, Al-Jamal and Ghadi (2008) showed that the high school leaving exam affected teachers' selection of teaching materials and techniques. Accordingly, those teachers adapted their teaching and focused mainly on preparing their students for this exam.

Arabic exams follow a traditional paper and pencil format with a specific focus on reading, writing, vocabulary, and grammar. Grammar is given special attention because of the grammar-centered teaching methods and also the complex nature of Arabic grammar. Interestingly, in most Arabic tests, there is a separate section assigned for poetry. This is not surprising given the historical connection between Arabs and poetry. Most of these exams usually use a combination of multiple choice and gap-filling formats, in addition to a writing section where students are asked to compose an essay or a paragraph depending on their grade level. Traditional test formats are preferred over alternative assessment because of their practicality. In most Arab countries, large classes are very common in schools, and consequently teachers attempt to use tests that are suitable for this context. Hence, it is very rare to find any testing of both speaking and listening skills. Before exams, teachers spend a considerable amount of time preparing their students using various materials and strategies. The preparation time increases substantially before final exams.

Until very recently, there were almost no standardized proficiency tests of Arabic as a native language. A very promising project for developing a proficiency test for native speakers of Arabic started recently at the United Arab Emirates (UAE) University (2010). The new test, which is called the Alain Test of Arabic Proficiency (ATAP), moves away from the grammar-oriented paradigm that is very common in academic circles in Arab countries and follows a communicative approach to language assessment. ATAP has four sections: reading, listening,

speaking, and writing. For both reading and listening a multiple choice format is used, while for writing and speaking test takers have to produce written and oral texts. The test is computer-based and it is scored automatically except for the writing and speaking sections. Since it is designed as a proficiency test, ATAP is not linked to any teaching context or specific curriculum. That is why it can be used in different contexts to assess the proficiency of native speakers in MSA. For example, the UAE Ministry of Education started using ATAP to assess the Arabic proficiency of their current teachers. Also, the UAE Ministry of Defense is planning to use ATAP with military personnel (Ali, personal communication, July 24, 2011). The UAE University aims also to promote the test in other Arab countries since there is no other competitor in the market to date.

### Evaluation

The review of the different Arabic language tests has shown a wide range of assessments used for different purposes in academic settings such as admission, placement, and exit decisions. In nonacademic contexts, these tests are mainly used for either hiring or promotion purposes, especially in AFL settings. Current trends in language assessment, which espouse new functions of assessment beyond gatekeeping, should be considered in Arabic language test development. For example, Chalhoub-Deville (2001) argues that the need for diagnostic testing is on the rise and consequently new diagnostic tests should be developed to address this issue. Another observation from this review is the very few tests in the first language context that measure general proficiency of Arabic. Adding to this, there is an urgent need for tests that target Arabic for specific purposes.

A unique problem associated with assessment of Arabic has to do with authenticity (i.e., the similarity between the test tasks and the target language use [TLU] domain). First of all, the diglossic situation in Arabic makes it extremely difficult to reach a definition of the native speaker. MSA is a variety which is only taught in schools and rarely used in everyday activities—except for formal contexts, such as journalism and news reporting. Accordingly, it is not easy to develop authentic assessment tasks that are reflective of the TLU domain. Elgibali and Taha (1995) analyzed the different tasks described in the ACTFL Arabic Proficiency Guidelines and concluded that listening and speaking tasks at both the advanced and intermediate levels require dialectal use, not MSA. For this reason and others, testing programs have recently developed new assessments for different dialects of Arabic (ACTFL, [www.languageTesting.com](http://www.languageTesting.com)).

The final remark in this discussion focuses on the psychometric qualities of the different L1/L2 Arabic tests. Testing programs involved in developing Arabic assessments need to provide more information about the reliability and validity of their test scores. Most of the testing programs do not offer sufficient data about the validation process in their test manuals. In addition, the field testing of AFL tests is problematic, since it is hard to find candidates with advanced Arabic proficiency as indicated by Winke and Aquil (2006). Furthermore, the investigation of score reliability in the OPI context does not go beyond checking inter-rater reliability. Although inter-rater reliability is a required procedure in this context, it is not sufficient. There is a clear need for using more sophisticated techniques, such



as generalizability theory and IRT, to look into the relative effects of different test facets on test scores (Gebril, 2009b, 2012). Finally, a number of Arabic tests reviewed in this chapter are used for making various decisions. However, little evidence is provided about the suitability of using test scores in making such decisions.

## Challenges and Future Directions

According to Al-Rajhi (2006), one of the biggest challenges facing Arabic language teaching is the absence of an academic body in charge of setting educational guidelines and standards. Schools independently develop their own set of acceptable standards guidelines and skills to be taught, which vary considerably from one institution to another. Many schools do not have a comprehensive articulation of standards, benchmarks, performance indicators, and instructional and assessment methods to be used. Classroom instruction around the Arab world remains largely textbook-based and teacher-centered, and standards have not found their way yet into classrooms, especially in public schools (Taha-Thomure, 2008). Arab countries taking the 2006 PIRLS test (Progress in International Reading Literacy) ranked 42, 43 and 44 out of 45 (PIRLS, 2006).

For Arabic as a foreign language, the same concerns prevail, except that ACTFL has developed standards for teaching Arabic as a foreign language. The standards are referred to as the 5Cs and include: cultures, communication, communities, connections, and comparisons. The difficulty, however, lies in regulating what the various Arabic language programs offer and moving to a more standards-based approach that aligns all aspects of the curriculum (Taha-Thomure & Lyman-Hager, 2009).

Standardized tests for Arabic as a foreign language have blossomed since the year 2000, and the field now has several high and low stakes tests to use including many that were outlined earlier in this chapter. However, for Arabic as a first language those tests remain scarce and unrelated to any national Arabic language arts standards and benchmarks (Sakr, 2008). Stakeholders will very soon need to call on assessment experts to help them define their purpose for having national Arabic literacy assessment and develop the performance and content standards it is based on.

Based on the previous discussion, it is clear that more attention should be paid to test fairness and ethics issues. These concepts should be regularly discussed among school administrators, test developers, teachers, and, more importantly students—particularly in the Middle East where there is little awareness about test fairness. There is also an urgent need for more transparency in test development and validation. Since Arabic tests are increasingly used in a number of contexts to make various decisions, validation studies are needed to check the appropriateness of these decisions. In addition, students should be provided with more information about their rights and responsibilities. Furthermore, the public should have access to different sources of information about how Arabic tests are validated, for what purposes they should be used, and what procedures are employed to ensure the accuracy of test scores. Such a course of action will help achieve the professionalization of the field of Arabic language assessment.