EXPERT AND INSTRUCTIONAL SYSTEM OF ARABIC LANGUAGE
FROM AN EDUCATIONAL PERSPECTIVE:
A COMPLEMENTARY APPROACH TO THE LINGUISTIC AND TECHNICAL
REPORTS of

M. Bawab, M. Mrayati, Y.M. Alam & M. H. Tayyan
The Higher Institute for Applied Sciences and Technology
Scientific Studies & Research Center
P.O. Box 4470 Damascus, Syria

by
Nimat Hafez Barazangi, Visiting Fellow
Women's Studies Program, 391 Uris Hall
Cornell University
Ithaca, New York 14853 USA

This paper is intended to summarize and complement the two reports prepared by Bawab et al., the linguistic and technical team from The Higher Institute for Applied Sciences and Technology (HIAST) headed by Dr. Muhammad Mrayati of the Scientific Studies and Research Center (SSRC) in Damascus, Syria. The first, entitled "PC-Based Conjugation of Arabic Verbs," was presented at the Arab School of Science and Technology's Second Spring Session on Information Technology and Applications, sponsored by the United Nations Industrial Development Organization and SSRC in May 25-31, 1992. The second (in Arabic), entitled "*al Nidham al Sarfi al Nahawi li al ⁹Arabiyah bi al Hasib* (Computerized Conjugational and Syntactical System of Arabic)," was presented at the Second Conference on Arabic Computational Linguistics in Kuwait, November 26-29, 1989.

The objective of this paper is twofold. One, to provide some suggestions concerning the educational applications of the Expert System (*Nidham Khabir*) and one of its sub-program, the PC-Based Conjugation of Arabic verbs . During my visit to the SSRC in January 1993, I was asked by the above team to develop these suggestions, based on my practical introduction to the system in addition to reading the above two reports.   Both, through the practical demonstration of the system by the team following prior discussions of the linguistic groundwork and information technology that underlie the system, and throughout the reading of the reports, I felt a need to make this work known to all users of Arabic. Thus, the second objective is to share this significant achievement of the Syrian team with Arabic linguists, Teachers of Arabic and scholars of Arabic and Islamic studies.

Introducing the System

The entire system is called "Expert System of Arabic (*Nidham Khabir*)" and is intended as part of a knowledge-based system of the Arabic language. Vax-11 was used to prepare the system, part of which is the PC-Based Conjugation of Arabic verbs that was "accomplished using IBM/PC/AT computer, and Turbo Prolog version 2 as a programming language." [1] The significance of the entire system is that its "output is characterized by thorough diacritics" and immediate and multiple applications such as: "language understanding, machine-aided translation, teaching, detection and correction of errors in Arabic word-processing systems." Additional significant applications of the program, from educational perspective, will be listed in a later section of this paper.

The system, like classical and standard Arabic dictionaries and other computerized systems of Arabic such as Bishai's Computerized Arabic Morphology, relies on the root (lexical entry) of the Arabic word as the input. Unlike the Bishai's Morphology program that is still not connected to Bishai's Computer Dictionary of Literary Arabic, The Expert System, through its production (*Ishtiqaq*) part, triggers the dictionary search that results in a display of the syntactical and conjugational rules from the most general, such as listing of augmented and unaugmented verbs, to the most specific, such as the combined conjugation rules of the *Hamza* (~) letter and *idgham* (gemination or doubling). The system is also distinguished, through its analytic (*tahlil*) part, by its ability to cue the user who does not have knowledge of Arabic roots into the different possible roots of any three-letter word or four-letter word that he or she may input.

The two main parts of the system, the Derivative (*Ishtiqaq* or generation and production of vocabulary) and the analytic (*Tahlil* or retrieval of the root with its vocalization, prefixes and suffixes) of the Arabic word, operate independently as well as interactively. Each part guides the user through an elaborate multiple-box and multi-window sub-system, from defining the original letters of a root (*Maaddah* or entry) in its different patterns (*Awzan*) and forms (*abwab*) to specifying all grammatical and morphological forms and derivatives.

The system is comprehensive in that it uses entries from all available Arabic-Arabic dictionaries, while Bishai's Dictionary relies on <u>al-Muhit al-Muhit</u> only. And, yet, the system is very specific and accurate in that it has over twenty-one thousand entries of augmented, unaugmented verbs in different combinations of voices, tenses, vocalization, etc.., while Bishai's Morphology Program has 5,000 verbs only, even though the Bishai's Dictionary has 50,000 entries. It was intended as a man-machine communication system of Arabic, both in its spoken and written forms. Thus, it provides lexical, morphological, syntactical and phonological analysis, synthesis and recognition of Arabic words. Furthermore, it provides special entries of newly introduced Arabic lexical entries for up-to-date scientific and engineering concepts.

In summary, this system combines morphological system with information technology systems, directed for the general user of Arabic whether to compose, to understand or to parse or produce different syntactical and grammatical structures, as well as for a scientist or an engineer, learning Arabic, or translating from and to the Arabic language.

Comments and Clarifications

The comprehensiveness of the system is very significant and opens many possibilities for future applications in understanding the language, for language development, acquisition, translation from and to Arabic, and dictionary spell-check as well as for discovering and correcting grammatical and syntactic mistakes.

I would like to draw the attention of the team to some weaknesses in the system from an educational point of view and how they may be modified into strengths of the system.

1. Introducing beginners to the system through the entire structural forms of a root (first window stating the different patterns and forms or *Awzan* and *Abwab*) in 18 different boxes might overwhelm the user or heighten his or her curiosity to the point of looking into all possible patterns of a root, and eventually getting lost. Developing a mechanism (such as contextual examples of the root in a maximum of three patterns or forms at a time) to introduce the learner to the appropriate pattern, according to his or her level of knowledge and the intended meaning may overcome this shortcoming. This mechanism, if developed, enables the learner to move either within a particular pattern and its derivatives derivatives, or within the entire system concurrent with his or her need and purpose.

2. Developing a semantic structure part for the system , where different possible meanings of each root are produced, may serve the above mechanism as well as the entire system in a more efficient process of derivation and analysis. That is, when a box of a particular pattern appears on the screen, a comparable box with the possible variation of meaning of the same root within the same pattern, but with different vocalization (*tanween*), embedded in a sentence will facilitate both understanding of the different grammatical, syntactic, and semantic rules as well as help the user assess and discover his or her mistakes.

3. It is not clear how the two parts, the *Ishtiqaq* and *tahlil*, are interconnected. Can a user move back and forth between them? or are they only accessible mechanically to produce the needed derivative or analysis? It will be very helpful for instructional purposes if the mechanical steps of derivation and analysis are made accessible to the user (i.e., appear on the screen, step-by step, as to how, when, and where, for example, a root acquires its prefixes, suffixes and

vocalization). This process will serve as a major teaching tool if designed for different levels of difficulty and complexity. Since the system uses color codes to distinguish the root from its various patterns, and the latters from the different conjugations, etc., other colors or box configurations and shapes may be also used to indicate different directions of derivation or analysis. Voice mechanism for simple, commonly used entries can also be added to instruct the young user and/or the adult non-native as well.

4. The entire system needs a help window which should explain the meanings, with examples, of patterns, forms, derivatives, etc.. This can be facilitated by using <u>one root</u> only and showing all the changes that it may encore within the different patterns, their vocalizations, etc., step-by-step. Drills that allow the user to parse or produce numerous and various forms of the entered root may facilitate this process further. This process also allows the user <u>to discover on his or her own</u> all the options and the limitation of a particular root. The user may also be able to retrieve a root from a particular form by analyzing a verb or a noun into its different parts.

5. The use of literary examples (i.e., a verb, a noun or a pronoun in a sentence) is essential to instruct the user into the different and possible semantic variations of the same root. This is particularly helpful in the use of prefixes and suffixes vs. similar letters that are part of the root that has more than three letters (e.g., it is not clear why the verb ^ʕ*Asawa* [p 12 of the Arabic report] maintained the *alif* after the *waw* in its third person plural past tense? What makes it different from *lam ya^ʕswa* in the apocopated (*Mabniy* or indeclinable) present tense, for instance? The same is true for other analytic examples, where printing the apocopating or indeclining pronoun to indicate the variations is essential for the user to make a distinction between prefixes and letters that belong to the original root (e.g., [p.14] *wa^ʕadat* is: either *w +^ʕAda +t* or *wa^ʕada +t)* but what are the semantic differences?

6. Disinential infliction (^ʕ*irab*)calls for more details in the derivative changes of a verb with the different inflictions. Explanations, with examples, of how a verb may take the same suffix but has different meanings with different inflictions (e.g. *Tardhayn* could be addressed to a second person singular or a second person plural, where its ^ʕ*irab* varies). How is the user expected to understand and distinguish the difference here?

Educational Applications of the System

From an educational perspective, three modules can be drawn. They are intended to facilitate the Arabic language acquisition to both native and non-native speakers through learner's experimentation and practice window. By installing within the system a user-friendly working window, and by developing a mechanism to show the user the reverse order of conjugation rules from the

conjugation table to the lexical entry (.i.e. how the particular word had evolved from its root), comprehensive instructional programs may be developed. An important aspect of these modules is that they can be broken down at different steps of complexity from introductory to intermediate and advanced levels of the grammatical structures, depending on the nature and ability of the user. To achieve this goal, the following steps may be taken within each module:

1. <u>First Module: An introductory level for a native speaker of Arabic who is either 3-5 year-old child who or a adult who is computer illiterate or unfamiliar with the Arabic grammar</u>

      Arabic as a first language is introduced in the spoken dialect (using the voice device for pronunciation) with picture-letter association. Touch screen mechanism may be used in place of the mouse or the keyboard to facilitate child play-like learning process. Gradual integration of the Modern Standard Arabic takes place through story telling, video games, and computer-assisted video instructions. This step is facilitated by instructions on the use of the mouse or the keyboard as well as the button or cue system into the multi- box, multi-window program of Arabic root and its simple and most commonly used derivatives.

2. <u>Second module: An introductory level for a child or an adult who is non-native speaker of Arabic</u>

   Arabic as a second/foreign language is introduced through formal/informal educational setting in an interactive culturally-oriented medium. Modern Standard Arabic and its grammatical structure are woven progressively in the content of other subjects as well.   This step is facilitated by voice-interactive and word-association programs that lead the user into the multi-box, multi-window program of Arabic root and its simple derivatives.

3. <u>Third Module: An intermediate and/or Advanced level for a learned[2]user</u>

      Arabic grammar is introduced in this module in a literary form through the teaching of different sentence structure and the analysis of texts from different subjects. The same types and modes of material used in the first and second modules can be modified to fit within this module, with the exception that the user may move directly into the next level of complexity of the lexical and morphological structure.

Conclusion

These proposed conceptual and linguistic integrative curricular modules operate within the theoretical and empirical findings of child language acquisition, language teaching, and bilingual education.   The universal principles of

language acquisition may assist in constructing the teaching of Arabic in relation to other languages. Conceptual learning is based on psycholinguistic and other developmental factors.   Contrary to the present methods of Arabic instruction and of the limited orientation of available materials, the proposed approach will enhance competency in all forms of the language.   Writing, reading texts, listening to and speaking in standard form are functions of well-designed curricular and instructional plans (as the experiment of Dr. Abdullah al-Dannan suggests)[3].   These plans dictate the nature and type of material.   In addition, instructional material should produce some level of coherence with the society's predominant language and culture in order to support children's cognitive development and the adult learners' intellectual needs.   In line with the traditional methods of teaching Arabic--recitation of the Qur'an, memorization of poetry and Qur'anic verses, and chanting of children's songs and rhymes-- the proposed approach will maintain the sociolinguistic aspects of Arabic, particularly in the context of dialect-speaking and/or non-Arabic speaking environment without causing linguistic nor cultural, neither intellectual confusion. By creating and/or separating existing learning environments, such as the home vis-`a-vis  the school for children or the literary and scientific vis-`a-vis the social setting for adults, neither the syntactic nor the semantic structures of the two languages (i.e., standard vs. colloquial or Arabic vs. English) get confused. This ability to segregate, and yet move flexibly between the two languages produces clear, unmixed speech pattern and pronunciation in addition to clear conceptual understanding and use of the two languages.

The application of the proposed approach, therefore, calls on developing material during the process of lesson planning and through constant piloting and evaluation by all concerned; the linguistic and technical team, the teachers/parents or language instructors and the curricular designers.   The learner will also have some control and manipulation of the material, particularly the computer-assisted video instructional devices.   Meanwhile, available traditional material may be used selectively to assist in early stages of planning and application.

1. Introducing beginners to the system through the entire structural forms of a root (first window stating the different patterns and forms or *Awzan* and *Abwab*) in 18 different boxes might overwhelm the user or heighten his or her curiosity to the point of looking into all possible patterns of a root, and eventually getting lost. Developing a mechanism (such as contextual examples of the root in a maximum of three patterns or forms at a time) to introduce the learner to the appropriate pattern, according to his or her level of knowledge and the intended meaning may overcome this shortcoming. This mechanism, if developed, enables the learner to move either within a particular pattern and its derivatives derivatives, or within the entire system concurrent with his or her need and purpose.

   2. Developing a semantic structure part for the system , where different possible meanings of each root are produced, may serve the above mechanism

as well as the entire system in a more efficient process of derivation and analysis. That is, when a box of a particular pattern appears on the screen, a comparable box with the possible variation of meaning of the same root within the same pattern, but with different vocalization (*tanween*), embedded in a sentence will facilitate both understanding of the different grammatical, syntactic, and semantic rules as well as help the user assess and discover his or her mistakes.

   3. It is not clear how the two parts, the *Ishtiqaq* and *tahlil*, are interconnected. Can a user move back and forth between them? or are they only accessible mechanically to produce the needed derivative or analysis? It will be very helpful for instructional purposes if the mechanical steps of derivation and analysis are made accessible to the user (i.e., appear on the screen, step-by step, as to how, when, and where, for example, a root acquires its prefixes, suffixes and vocalization). This process will serve as a major teaching tool if designed for different levels of difficulty and complexity. Since the system uses color codes to distinguish the root from its various patterns, and the latters from the different conjugations, etc., other colors or box configurations and shapes may be also used to indicate different directions of derivation or analysis. Voice mechanism for simple, commonly used entries can also be added to instruct the young user and/or the adult non-native as well.

   4. The entire system needs a help window which should explain the meanings, with examples, of patterns, forms, derivatives, etc.. This can be facilitated by using <u>one root</u> only and showing all the changes that it may encore within the different patterns, their vocalizations, etc., step-by-step. Drills that allow the user to parse or produce numerous and various forms of the entered root may facilitate this process further. This process also allows the user <u>to discover on his or her own</u> all the options and the limitation of a particular root. The user may also be able to retrieve a root from a particular form by analyzing a verb or a noun into its different parts.

   5. The use of literary examples (i.e., a verb, a noun or a pronoun in a sentence) is essential to instruct the user into the different and possible semantic variations of the same root. This is particularly helpful in the use of prefixes and suffixes vs. similar letters that are part of the root that has more than three letters (e.g., it is not clear why the verb *ʕAsawa* [p 12 of the Arabic report] maintained the *alif* after the *waw* in its third person plural past tense? What makes it different from *lam yaʕswa* in the apocopated (*Mabniy* or indeclinable) present tense, for instance? The same is true for other analytic examples, where printing the apocopating or indeclining pronoun to indicate the variations is essential for the user to make a distinction between prefixes and letters that belong to the original root (e.g., [p.14] *waʕadat* is: either *w +ʕAda +t* or *waʕada +t)* but what are the semantic differences?

6. Disinential infliction (ᶜirab)calls for more details in the derivative changes of a verb with the different inflictions. Explanations, with examples, of how a verb may take the same suffix but has different meanings with different inflictions (e.g. *Tardhayn* could be addressed to a second person singular or a second person plural, where its ᶜirab varies). How is the user expected to understand and distinguish the difference here?

The objective of this paper is twofold. One, to provide some suggestions concerning the educational applications of the Expert System (*Nidham Khabir*) and one of its sub-program, the PC-Based Conjugation of Arabic verbs .

Notes

1. All quotations are from the above two reports, unless indicated otherwise.

2. "Learned" here indicates a level of knowledge that qualify the user to move directly into the third module without the need for introduction to the computer operational steps, nor to the primary stages of the system, neither to the elementary level of the Arabic language.

3.   Al-Dannan, Abdullah. "al ᶜibdaᶜ wa al Llughah al ᶜArabiyah fi al Manahij al Madrasiyah (Creativity and the Arabic Language in School Curricula)" Kuwait University, College of Education. No date or name of journal listed.

al-Ahram Center for Translation and Publishing 1987; and Ads 1980)

Though the East coast Consortium of Universities in the United States decisded that instruction in the summer institute at Midelbury would be in Modern Standard Arabic for all four skills (Allen 231-2), academic institutions in the United States and schools in the Arab world rarely teach speaking skills in standard or classical Arabic.

Defense Language Institute. 1986. **Master Paln for a Strategy of Excellence**. Monterey, CAL: Defense Language Institute.

both classical Arabic philology (cf. Khalil 1993) and the relation of Arabic dialects to language development (Abd Al-Tawwab 1967), and on modern linguistics and learning theories (Freudenstein, et. al. 1986).

Abd Al-Tawwab, Ramadan. 1967. Lahn alᶜAmmah wa al tatwur al lughawi (Popular dialects and Linguistics Development). (Cairo: Dar al Maᶜarif).

Freudenstein, Reinhold, James, c. Vaughan, Ed. 1986. Confidence Through Competence in Modern Language Learning. CILT Reports and Papers 25. Selection of Papers from the World Conference of the International Federation of

Modern Language Teachers in Helsinki, Finland, 1985. (London: Centre for Information on Language Teaching and Research).

Freudenstein, Reinhold, James, c. Vaughan, Ed. 1986. Confidence Through Competence in Modern Language Learning. CILT Reports and Papers 25. Selection of Papers from the World Conference of the International Federation of Modern Language Teachers in Helsinki, Finland, 1985. (London: Centre for Information on Language Teaching and Research).

Khalil, Hilmi. 1993. Muqaddima li Dirasat Fiqh al Lugha (Introduction to the Studies in Linguistics (Alexandria, Egypt: Dar al Ma^{ς}rifah).

 This integrated approach relies on the applications generated from knowledge of first and second language acquisition (Harley 1986, Kharma 1983), and applied linguistics (Hart 1987), and instruction by computer-based interactive video (Kenny, 1992).

 NOTE 1 These CALI programs are largely template-based or rely on videodiscs. Also, CALI programs either depend on dictionary for glosses, in which beginning students have difficulty determining the the entry form, or partial systems of Arabic morphology and dictionary, such as Bishai's 1992, that are not interconnected.   On the other hand, the Syrian system is comprehensive in its morphological and syntactical structures, and by being connected to a dictionary/lexical system at which a learner may access any root by mere entering a word in any form. Also, being a knowledge-based expert system , it can operate as a data-base from which several programs can be derived even by the learners to fit their particular needs.

2. These approaches vary in their view of language teaching (e.g. audiolingual, individualized, total immersion), in their view of Arabic among other foreign languages (its level of difficulty, amount of time needed to be mastered), and nature and goals of skills being emphasized (speaking- vs. reading-skills and professional competence vs achievement-based programs) by different institutions that teach Arabic (Allen 1992: 222-233).

3. wrote about Arabic morphological analyser generator (Thalouth and Al-Dannan 1990) and

5.   One of the goals set by the Defense Language Institute (1991) is to add a cultural background in the training of teachers of Foreign language and to grant an associate degree in the field to its instructors of "less commonly taught languages," Arabic included. Similar policies in academic institutions, represented mainly by ACTFL (Allen, 1990 and Dandonoli 1991) and NCOLCTL, to develop language specific Language Learning Frameworks, including Arabic (Parkinson 1993) are another example of recent attempts to improve the standards of Arabic teaching/learning as a Foreign Language. Parkinson (1992) also discusses the importance of computer-aided programs in enhancing language teaching in general. One of the early sessions in 1980 of the Arab School of Science and Technology was devoted to computer technology. Subsequent sessions (1983, 89, 90 and 91) were also devoted to computers, signals and information processing and applied Arabic linguistics. Furthermore, one of the school's basic goals is to encourage Arab scientists to return to [or connect with] their countries [or countries of origin] by providing them with opportunities to contribute to the school's activities and closely review the scientific resources of the respective country (Mackay 1990).

8. This aspect of program development could use knowledge and experience of previous work on Computer Assisted Language Learning (CALL), particularly the structural module developed by Ruschoff (1988: 107). In Ruschoff's structure, exercise modules are drawn from a knowledge base. In

turn, the learners input their "previous knowledge, learning proficiency, learning style, and performance" into their own module. "Tutorial component" is generated and is imputed into the exercise modules. I will elaborate on such a module by enhancing the communication circuit with the teacher input through interactive windows, and instead of having the total input feeding back only into the exercise modules, I would make the input-output interaction complete the circle further by including the knowledge base as well .

9. Khaled (1984: Abstract) discusses the use of verbs as lexical items and grammatical forms and the form-function relationship in teaching English for special purposes (ESP) in nursing. I further suggest that the method of using linguistic analysis to select appropriate instructional materials can best be utilized in the learning of the language under considerations as well, and in designing curricula for that purpose.

Al-Hamad, Fayiz 1983. "Problems in Teaching Arabic to Foreigners." In Proceedings of the Second Annual Linguistics Conference, edited by Jonathan Owens and Issam Abu Salim, 81-100. Irbid, Jordan: Yarmouk University.

Dandaloni, Patricia. 1991. "The ACTFL Computerized Adaptive Test of Foreign Language Reading Proficiency" in Hastings, O. Smith, WM. Flint, Ed. Modern Technology in Foreign Language Education" Applications and Projects. The ACTFL Foreign Language Education Series. (American Council on the Teaching of Foreign Languages).

Harley, Brigit. 1986. Age in Second Language Acquisition (San Diego: College-Hill Press).

Hart, Robert S. 1987 "Current Trends in Computer-Based language Instruction." IDEAL (2, special issue): 63-74.

Kenny, Richard F. 1992. "The Generative Effects of Instructional Organizers with Computer-Based Interactive Video." in Proceedings of Selected Research and Development Presentation at the Convention of the Association for Educational Communications and Technology.

Kharma, Nayef. 1983. A Constructive Analysis of the Use of Verb Forms in English and Arabic. Studies in Descriptive Linguistics, General Editor: Dietrich Nehls, Volume 10 (Heidelberg: Julius Groos Verlag).

Lambert, Richard, et. al. 1984. Beyond Growth: The Next Stage in Language and Area Studies.

Richardson, M.D.; and Others. "Computer Technology: Bridging the Gap." Paper presented at the Annual Meeting of the USA/China Education Consortium.

Roach, Peter. 1986. "Talking and Listening With A Computer." in Shivtiel, Avihai, Ed. Teaching Arabic with the Computer: Proceedings of the First Symposium held at the University of Leeds, July 21- 22 1986:9-16.

Thalouth, B. Al-Dannan, A. 1990. "A comprehensive Arabic Morphological Analyzer Generator." in Mackay, P.   Computers and the Arabic Language, 208-17.