Literacy Studies: Perspectives from Cognitive Neurosciences, Linguistics, Psychology and Education

Elinor Saiegh-Haddad
R. Malatesha Joshi *Editors*

Handbook of Arabic Literacy

Insights and Perspectives

Springer

Handbook of Arabic Literacy

LITERACY STUDIES

While language defines humanity, literacy defines civilization. Understandably, illiteracy or difficulties in acquiring literacy skills have become a major concern of our technological society. A conservative estimate of the prevalence of literacy problems would put the figure at more than a billion people in the world. Because of the seriousness of the problem, research in literacy acquisition and its breakdown is pursued with enormous vigor and persistence by experts from diverse backgrounds such as cognitive psychology, neuroscience, linguistics and education. This, of course, has resulted in a plethora of data, and consequently it has become difficult to integrate this abundance of information into a coherent body because of the artificial barriers that exist among different professional specialties. The purpose of this series is to bring together the available research studies into a coherent body of knowledge. Publications in this series are of interest to educators, clinicians and research scientists in the above-mentioned specialties. Some of the titles suitable for the Series are: fMRI, brain imaging techniques and reading skills, orthography and literacy; and research based techniques for improving decoding, vocabulary, spelling, and comprehension skills.

For further volumes:
http://www.springer.com/series/7206

Elinor Saiegh-Haddad • R. Malatesha Joshi
Editors

Handbook of Arabic Literacy

Insights and Perspectives

Volume 9

Springer

Editors
Elinor Saiegh-Haddad
Bar-Ilan University
Ramat Gan
Israel

R. Malatesha Joshi
Texas A&M University
College of Education
College Station
Texas
USA

ISSN ISSN 2214-000X
ISBN 978-94-017-8544-0
DOI 10.1007/978-94-017-8545-7
Springer Dordrecht Heidelberg New York London

ISSN 2214-0018 (electronic)
ISBN 978-94-017-8545-7 (eBook)

Library of Congress Control Number: 2014932265

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Foreword

Language defines humankind; writing defines civilization (Daniels 1996). Modern civilization was then redefined by the printing press and paper. And today, writing technology is again transforming our world as the electronic media turn the world on paper (Olson 1994) into a paperless world. Yet the key to the world of print, whether on a computer screen, paper, clay or stone remains unchanged; the near-instantaneous access to the meanings locked in the symbol strings of the written text. Around half of humanity, however, does not hold this key. The illiterate and semi-literate are excluded. In most parts of Asia and Africa, illiteracy and poverty go hand in hand. Curiously, in the Arabic-speaking world, literacy levels are uniformly and alarmingly low in wealthy and impoverished societies alike. Even highly educated and skilled readers of Arabic read their native Arabic more slowly than they read non-native languages such as English, Hindi or Arabic's Semitic cousin Hebrew which shares the same highly synthetic poly-morphemic structure as Arabic. Why is literacy learning so difficult in Arabic?

In addressing this quandary, the present volume offers no quick-fix remedies, but it does offer a first-generation infrastructure of scientific theory and research that can inform decision-making by policy-formulators, educators and practitioners confronting the literacy challenge in Arabic on a daily basis. Saiegh-Haddad and Joshi have rendered an outstanding service to the field in this ground-breaking volume which brings together a panoply of leading scholars from the Middle East, North America, and Europe, representing a wealth of disciplinary perspectives. The depth and breadth of the scholarship will no doubt earn this handbook benchmark status for future work in this field.

Arabic is the fourth most common language in the world, and the Arabic script is the second most widely used segmental (phonemic) script after Roman. The scholarship embodied in this volume will not only inform practitioners and researchers of the Arabic language and literacy but any theory aspiring beyond language-specific status. For too long, the language and literacy research agenda has been a captive of Anglo-American concerns, overwhelmingly dominated by English. Today, the world is finally waking up to the fact that most of the world's languages are not English-like. This Anglo-centrism is ever more poignant in the literacy domain given that English orthography is an outlier even among European alphabets (Share

2008). Most of the world's literacy learners are learning to read in languages and writing systems that are neither alphabetic (i.e., full and equal status is given to consonant and vowel signs) nor European. It therefore behooves researchers to keep informed of literacy studies across a range of languages and writing systems and avoid the scientific solipsism of the past in which literacy studies in languages other than English were regarded as mere exotic *curiosa*. This volume will become a landmark not merely because it is a world first, but because it offers *all* literacy scholars a wider angle lens on their own work.

September, 2013 David L. Share
Haifa University, Israel

Preface

It is customary in promoting a book to talk about an existing lacuna in the field to which the book belongs. This may or may not in fact be the case, but this tradition does indeed match the reality in the field of Arabic linguistic studies as far as this volume is concerned. Studies of Arabic literacy are meagre and they remain marginal in Arabic linguistics, in spite of the undisputed importance of this topic in understanding the language at the crossroads of psycholinguistics and language acquisition, educational linguistics, sociolinguistics and cultural studies. The multifaceted nature of this topic is reflected in the content of this volume of essays which report the findings of new research, or bring together the major insights of existing work to map aspects of Arabic literacy studies for use as a platform for future research. The net result is a volume of great reach, depth and interest. It describes, explains and offers empirical and quantitative conclusions which can help interested scholars reflect, comparatively, on literacy in Arabic and other languages from theoretically-informed perspectives.

In recent years, Arabic literacy has emerged as an issue of great educational importance in the Arabic speaking world. PIRLS results during the past few years have consistently placed participating Arab countries at the bottom of international achievement levels. Arab policy-makers and pedagogy experts have been exercised by this and are on the look-out for ways to understand the problem and to devise solutions. Arabic language teaching reforms in Arab countries during the last decade are an expression of this endeavour (I know this to be the case from my long experience in this field). Although the essays in this volume are not offered as a solution to this problem, they nevertheless provide a basis from which an understanding of it can be developed. This understanding is bound to be complex and may speak in different inflections, depending on disciplinary perspective.

This is an excellent volume and the first of its kind. It will be the first port of call for those who wish to learn about Arabic literacy. The editors and contributors are to be congratulated on this achievement.

April 2013

King's College
Cambridge University
U.K.

Yasir Suleiman

Introduction

Among the various reasons for literacy problems that have been postulated, Vellutino et al. (2003) cite instruction and environment as being the two most fundamental factors. Instructional factors include the lack of a suitable literacy environment in schools, ineffective instructional methods, and the teachers' lack of knowledge about language and structure (Cunningham et al. 2004; Joshi et al. 2009; McCutchen et al. 2002; Moats and Foorman 2003; Piasta et al. 2009). Environmental reasons include poor oral language development (Piasta and Wagner 2010), number of books available at home, parental attitudes, and parental models (Chiu and McBride-Chang 2006).

In addition to these factors, orthography may also influence literacy acquisition. In a seminal study, Seymour et al. (2003), examined word reading of children in grades 1 and 2 in 13 European orthographies and found that children who were learning to read in transparent orthographies such as Finnish, German, and Spanish read words faster and more accurately than children who were learning to read in opaque orthographies such as English and French. However, the majority of the studies conducted on literacy acquisition have been conducted on children speaking English, which, according to Share (2008), is an 'outlier' orthography.

There are very few studies on literacy acquisition among speakers of Arabic, even though it is the fourth most spoken language in the world. Further, Arabic orthography depicts interesting linguistic and orthographic features and hence offers an excellent testing ground for various competing theories of language and reading acquisition. These features include diglossia, double-script, vowelization/vocalization, root-based morphological structure and morpho-syntactic marking, to mention a few.

The chapters included in this book address linguistic, orthographic, cognitive, as well as environmental and socio-cultural factors in literacy development in Arabic. Besides being the first edited book of empirical research into language and literacy development in Arabic, it provides a representation of recent approaches to the study of Arabic literacy as well as a demonstration of the theoretical models, methods, and tools that have been recently employed in addressing literacy-related questions in Arabic. The handbook brings together a range of perspectives on the topic of literacy acquisition in Arabic and offers a discussion of the theoretical frameworks as well as the practical implications of the questions investigated. Rather than provide definitive

answers to questions regarding processing, instruction or intervention, the aim of the handbook is to offer a synthesis of contemporary research insights and perspectives on the study of Arabic literacy in the hope of generating more research interest in a hitherto neglected area of investigation. Here, we would like to thank the contributors as well as the anonymous reviewers for their invaluable contribution to this project.

Because an understanding of literacy development in any language requires first and foremost an accurate and explicit understanding of the linguistic and orthographic structure of that language, the handbook opens with an introductory descriptive chapter, co-authored by Saiegh-Haddad and Henkin-Roitfarb, that provides an outline of the structure of Arabic language and orthography with specific focus on aspects of Arabic linguistic structure that have direct implications for literacy development. The chapter provides a linguistic description of Arabic, yet care was taken to ensure that its content is accessible to readers with no background in Linguistics or knowledge of the Arabic language.

The remaining chapters in this collection are clustered into five thematic parts. Part two focuses on morphological structure and orthographic complexity and features psycholinguistic research into the representation and processing of Arabic words—how information moves from the page into the lexicon of the readers—and it includes four chapters. Chapter 2, by Boudelaa, addresses the nature of the Arabic lexicon and uses evidence from spoken and written word recognition in order to probe whether the Arabic mental lexicon is morpheme-based or stem-based. In Chap. 3, Funder-Hansen addresses word recognition in root-based Arabic and uses the unique features of Arabic script and Semitic morphology to propose a language-specific model of reading. Chapter 4 also addresses orthographic features in word reading in Arabic. The authors, Eviatar and Ibrahim, synthesize the insights they have gained from a series of recent examinations of word reading in Arabic and discuss the factors that they believe contribute to difficulty in developing this ability.

Part three focuses on reading and spelling development and disorders in Arabic. In Chap. 5, Mohamed, Landerl and Elbert report an epidemiological survey of specific reading and spelling disorders in Arabic speaking children in Egypt. This study reveals a less than expected dissociation between reading and spelling in vowelized Arabic compared to other shallow orthographies, as well as a high incidence of specific reading and spelling disorders in Arabic speaking children in Egypt. In Chap. 6, Friedmann and Haddad-Hanna discuss evidence demonstrating various types of developmental dyslexias in Arabic and present new research directions that utilize orthographic features of Arabic in understanding reading breakdown. In Chap. 7, Ravid, Naoum and Nasser report a study of narrative text production in Arabic in an attempt to shed light on the developing language basis of literacy. Abu Ahmad, Ibrahim and Share report a longitudinal study from kindergarten to grade 2 of the cognitive predictors of early reading ability in Arabic in Chap. 8. Using modularity as a framework, they show that while early word recognition depends primarily on phonological abilities, reading comprehension still relies heavily on decoding as well as higher-order linguistic and cognitive abilities.

Part four, which contains five chapters, addresses various aspects of Arabic diglossia. In Chap. 9, Myhill reports comparative data on literacy rates in a number

of countries and shows that basic literacy rates in Arabic-speaking countries are far lower than would be expected based upon their relative wealth. Using comparative evidence, he argues that much of the explanation for this lies in their usage of a standard language which is based upon an earlier version of the language which no one speaks anymore, and that the best policy for addressing this problem in initial literacy instruction would appear to be to use a strategy parallel to that adopted for languages such as Chinese, Japanese, and Sinhala in which early literacy is based on written phonological representations of the different spoken dialects. In the wake of this latter proposal, in Chap. 10, Saiegh-Haddad and Spolsky discuss some of the problems, ideological and others, in basing initial literacy in a diglossic context on the spoken vernacular. Then, the authors describe a pioneering attempt to address these problems in literacy development in Arabic. Chapter 11, authored by Laks and Berman, describes a novel approach to studying the linguistic manifestation of diglossia by analysing the linguistic structure of oral and written narrative text productions in spoken and standard Arabic, respectively, by Jordanian native speakers.This examination qualifies the linguistic distance between spoken Arabic and standard Arabic as reflected in the actual use of the two language varieties in oral and written text production. In Chap. 12, Rosenhouse examines another reflection of diglossia in the language used in textbooks in Israeli Arabic-speaking schools. The study analyses the language used in the textbooks and its proximity/distance from the language of speakers in an attempt to gain insight into the consistency, or lack thereof, in the linguistic elements that are covered in these textbooks, as well as of the suitability of the texts to the young learners and their effectiveness in promoting language acquisition. Chapter 13, authored by Khamis-Dakwar and Makhoul describes the rationale and research evidence behind the construction of a novel language assessment tool—ADAT (Arabic Diglossic knowledge and Awareness Test) that aims at measuring diglossic knowledge development in typically developing native Arabic-speaking children.

Part five addresses socio-cultural aspects of literacy development in Arabic. Chapter 14, authored by Tibi and McLeod, reports a study of the acquisition of emergent literacy in the Emirate of Abu Dhabi in the United Arab Emirates. In particular, it examines the language and literacy acquisition consequences of a newly implemented bilingual educational plan in the country—the "New School Model", which entails bilingual education (Arabic & English) from kindergarten through the years of compulsory schooling. Chapter 15, authored by Korat, Aram, Hassunha-Arafat, Hag-YehiyaIraki, and Saiegh-Haddad, is a study of the quality of storybook reading and joint word writing by Arabic speaking mothers with their young children. The study tested the influence of these activities, as well as socio-economic status and home literacy environment, on children's literacy attainment and provided insights into the design of family intervention programs so as to maximize children's literacy growth within the Arabic-speaking family.

Part six includes three chapters that address literacy development in special populations. These populations include bilingual English-Arabic speakers in the U.S.A., Arabic foreign language learners in Israel, and Braille reading of Arabic native speaking blind individuals. Chapter 16, authored by Farran, Bingham and

Table 1 Index to the symbols used in the transcription of Arabic vowels

Sound symbol	Sound description	Arabic grapheme and name
/*a*/	short low vowel	فتحة: ◌َ
/*a:*/	long low vowel	الف ا
/*u*/	short high back vowel	ضمة: ◌ُ
/*u:*/	long high back vowel	واو: و
/*i*/	short high front vowel	كسرة: ◌ِ
/*i:*/	long high front vowel	ياء ي

Matthews, reports a study of the role of environmental variables (parent education, beliefs, and home language use and literacy practices) in language and literacy outcomes among English-Arabic bilingual children in the US, and reveals a strong relationship between parent home language use and the development of various language and literacy skills in Arabic in this population. Chapter 17 describes two studies of the acquisition of grapho-phonemic representations among native Hebrew speakers learning Arabic as a foreign language. Based on quantitative and qualitative analyses of spelling errors among eighth graders during the second year of exposure to the written form of Arabic, and an examination of the developmental trajectory of grapho-phonemic knowledge among eighth, ninth, and tenth graders, Russak and Fragman demonstrate slow progress in spelling accuracy in this population and suggest that the phonological distance between Arabic and Hebrew may be one important cause. The last chapter in this collection, Chap. 18, authored by Jarjoura and Karni is unique in testing Braille reading in blind and sighted Arabic native speakers. The study reports the findings from Braille reading tasks of vowelized and unvowelized words and texts in Arabic. It shows, *inter alia*, that Arabic Braille readers, children and adults, are pervasively slower compared to English Braille readers. On the basis of these results, as well as the analysis of errors, the authors argue that specific characteristics of Arabic, including diglossia and vowelization may be responsible for the observed slowness in Braille reading.

Transcription Conventions

All chapters included in this collection follow uniform phonemic transcription and indexing conventions. The transcription of Arabic words follows a broad phonemic transcription system, unless in cases where a phonetic transcription was required. The phonetic symbols used are a combined modified version of the IPA (International Phonetic Alphabet) and the APA system used by American linguists. An index to the phonetic symbols used in representing Arabic sounds is provided in the tables below (Tables 1 and 2). Slant lines are used to enclose phonemes presented in an italicized font (e.g., /*b*/, /*m*/). No slant lines are used to enclose the transcrip-

Table 2 Arabic consonants and corresponding Arabic letters

		Stop	Fricative	Affricate	Liquid	Nasal	Glide
Labial		/b/ ب				/m/ م	/w/ و
Labiodental			/f/ ف				
Dental	plain	/t/ ت /d/ د	/s/ س /z/ ز		/l/ ل		
	emphatic	/ṭ/ ط /ḍ/ ض	/ṣ/ ص				
Interdental	plain		/θ/ ث /ð/ ذ				
	emphatic		/ðˁ/ ظ				
Alveolar			/š/ ش /j/ ج	(/ž/) ج	/r/ ر	/n/ ن	
Palatal							/y/ ي
Velar		/k/ ك	/x/ خ /ɣ/ غ				
Uvular		/q/ ق					
Pharyngeal			/ħ/ ح /ʕ/ ع				
Glottal		/ʔ/ ء	/h/ ه				

tion of full words, however (e.g., *walad*). Square brackets [ ] are used for phonetic transcription and quotes are used for English glossing (e.g., *walad* 'boy'); where necessary, the actual Arabic word is also provided. A hyphen—is used to mark morpheme boundaries (e.g., *l-walad* 'the boy'; *bi-bayt-i* 'in my house') and dots are used to mark syllable boundaries (e.g., *mak.ta.bu.na:* 'our desk'). Where internal morphological structure is relevant, capital letters are used for root consonants (e.g., KTB) and capital C for the consonant slots of word patterns (e.g., CaCaCa). Capital letters are also used to represent the letters of written words, (e.g., KTB, KATB, MKTUB).

References

Bauer, T. (1996). Arabic writing. In P. T. Daniels & W. Bright (Eds.), *The world's writing systems* (pp. 559–746). New York: Oxford University Press, Inc.

Chiu, M., & McBride-Chang, C. (2006). Gender, context, and reading: A comparison of students in 43 countries. *Scientific Studies of Reading, 10,* 331–362.

Coulmas, F. (1996). *The Blackwell encyclopedia of writing systems*. Malden: The Blackwell Publishers, Ltd.

Cunningham, A. E., Perry, K. E., Stanovich, K. E., & Stanovich, P. J. (2004). Disciplinary knowledge of K–3 teachers and their knowledge calibration in the domain of early literacy. *Annals of Dyslexia, 54,* 139–167.

Henderson, L. (1982). *Orthography and word recognition in reading*. London: Academic Press.

Joshi, R. M., Binks, E., Hougen, M., Dahlgren, M., Dean, E., & Smith, D. (2009). Why elementary teachers might be inadequately prepared to teach reading. *Journal of Learning Disabilities, 42,* 392–402.

Lewis, M. P. (Ed.). (2009). *Ethnologue: Languages of the world* (16th ed.). Dallas: SIL International.

McCutchen, D., Abbott, R. D., Green, L. B., Beretvas, S. N., Cox, S., Potter, N. S., et al. (2002). Beginning literacy: Links among teacher knowledge, teacher practice, and student learning. *Journal of Learning Disabilities, 35,* 69–86.

Moats, L. C., & Foorman, B. R. (2003). Measuring teachers' content knowledge of language and reading. *Annals of Dyslexia, 53,* 23–45.

Piasta, S. B., Connor, C. M., Fishman, B., & Morrison, F. J. (2009). Teachers' knowledge of literacy concepts, classroom practices, and student reading growth. *Scientific Studies of Reading, 13,* 224–248.

Piasta, S. B., & Wagner, R. K. (2010). Developing early literacy skills: A meta-analysis of alphabet learning and instruction. *Reading Research Quarterly, 45,* 8–38.

Seymour, P. H. K., Aro, M., & Erskine, J. M. (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology, 94,* 143–174.

Share D. L. (2008). On the anglocentricities of current reading research and practice: The perils of overreliance on an "outlier" orthography. *Psychological Bulletin, 134,* 584–615.

Vellutino, F. R., Scanlon, D. M., & Jaccard, J. (2003). Toward distinguishing between cognitive and experiential deficits as primary sources of difficulty in learning to read: A two-year follow-up of difficult to remediate and readily remediated poor readers. In B. R. Foorman (Ed.), *Preventing and remediating reading difficulties: Bringing science to scale* (pp. 73–120). Baltimore: York.

Contents

Contributors

Hanadi Abu Ahmad Learning Disabilities Department, University of Haifa, Israel

Dorit Aram Tel Aviv University, Tel Aviv, Israel

Ruth A. Berman Department of Linguistics, Tel Aviv University, Ramat Aviv, Israel

Gary E. Bingham The College of Education, Georgia State University, Atlanta, Georgia, USA

Sami Boudelaa Department of Linguistics, United Arab Emirates University, Al Ain, United Arab Emirates

Neuroscience of Language Laboratory, NYU Abu Dhabi, Abu Dhabi, United Arab Emirates

Thomas Elbert Department of Psychology, University of Konstanz, Konstanz, Germany

Zohar Eviatar Psychology Department, University of Haifa, Haifa, Israel

Lama K. Farran The College of Education, Georgia State University, Atlanta, Georgia, USA

Alon Fragman Ben Gurion University, Be'er Sheva, Israel

Naama Friedmann Language and Brain Lab, School of Education and Sagol School of Neuroscience, Tel Aviv University, Tel Aviv, Israel

Gunna Funder Hansen (Dep.) Svendborg Langeland, HF & VUC FYN (Adult Education Centre Fyn), Svendborg, Denmark

Manar Haddad-Hanna Tel Aviv University, Tel Aviv, Israel

Himat Hag-Yehiya Iraki School of Education, Bar-Ilan University, Ramat Gan, Israel

Safieh Hassunha-Arafat School of Education, Bar-Ilan University, Ramat Gan, Israel

Roni Henkin-Roitfarb Ben-Gurion University of the Negev, Beersheba, Israel

Raphiq Ibrahim Learning Disabilities Department, University of Haifa, Israel

The Edmond J. Safra Brain Research Center for the Study of Learning Disabilities, University of Haifa, Haifa, Israrel

Waleed Jarjoura The Edmond J. Safra Brain Research Center for the Study of Learning Disabilities, University of Haifa, Haifa, Israel

The "Convent of Nazareth" school for the blind, Nazareth, Israel

Avi Karni The Edmond J. Safra Brain Research Center for the Study of Learning Disabilities, University of Haifa, Haifa, Israel

The Sagol Department of Neurobiology, University of Haifa, Haifa, Israel

Reem Khamis-Dakwar Department of Communication Sciences and Disorders, Adelphi University, Garden City, NY, USA

Ofra Korat School of Education, Bar-Ilan University, Ramat Gan, Israel

Lior Laks Department of English Linguistics and Literature, Bar-Ilan University, Ramat Gan, Israel

Karin Landerl Department of Psychology, University of Graz, Graz, Austria

Baha Makhoul Haifa University, Haifa, Israel

The Centre for Educational Technology (CET), Tel-Aviv, Israel

Mona W. Matthews The College of Education, Georgia State University, Atlanta, Georgia, USA

Lorraine McLeod UAE University, Al Ain, United Arab Emirates

Wessam Mohamed Department of Educational Psychology, Faculty of Education, Fayoum University, Faiyum, Egypt

John Myhill Deptartment of English Language and Literature, University of Haifa, Haifa, Israel

Dina Naoum The Department of Communications Disorders, Tel Aviv University, Tel Aviv, Israel

Suheir Nasser The Department of Communications Disorders, Tel Aviv University, Tel Aviv, Israel

Dorit Ravid School of Education and the Department of Communications Disorders, Tel Aviv University,, Tel Aviv, Israel

Judith Rosenhouse Swantech Ltd., Haifa, Israel

Susie Russak Beit Berl Academic College, Kfar Saba, Israel

Elinor Saiegh-Haddad Department of English Linguistics and Literature, Bar-Ilan University, Israel

Department of English Linguistics and Literature, Bar-Ilan University, Ramat Gan, Israel

David L. Share Learning Disabilities Department, University of Haifa, Israel

The Edmond J. Safra Brain Research Center for the Study of Learning Disabilities, University of Haifa, Haifa, Israrel

Bernard Spolsky Bar-Ilan University, Ramat Gan, Israel

Sana Tibi Queen's University, Kingston, ON, Canada

About the Authors

Elinor Saiegh-Haddad is a professor of Linguistics in the English Department of Bar-Ilan University, Israel and a consultant for the National Authority for Testing and Evaluation of the Israel Ministry of Education (RAMA) as well as the Centre for Educational Technology (CET). Her main research areas include the acquisition of reading and writing in bilingual children, and in Arabic diglossia, and the relationship between oral language skills and the acquisition of literacy. Prof. Saiegh-Haddad has published numerous articles on literacy acquisition in children and co-edited (with Esther Geva) a special issue of Reading and writing on the simultaneous acquisition of reading in two languages. She has also been active in curriculum reform and development in Israel in the field of literacy (Arabic as L1 and English as FL) at both the elementary and pre-school levels.

R. Malatesha Joshi Ph.D. is Professor of Reading/Language Arts Education, ESL and Educational Psychology at Texas A&M University, where he teaches and conducts research in literacy development and literacy problems among monolinguals and bilinguals. Dr. Joshi is the founding editor of *Reading and Writing: An Interdisciplinary Journal*, rated one of the top journals in education and educational research. A monograph series titled *Literacy Studies: Perspectives from Cognitive Neurosciences, Linguistics, Psychology and Education* is also under his editorship.

Part I
The Arabic Language

Chapter 1
The Structure of Arabic Language and Orthography

Elinor Saiegh-Haddad and Roni Henkin-Roitfarb

Abstract This chapter was designed to promote our understanding of the triangulation, in Arabic, of language, orthography and reading. We focus on topics in the structure of the Arabic language and orthography that pertain to literacy research and practice. It is agreed that the development of basic reading skills is influenced by linguistic (mainly phonological and morpho-syntactic) and orthographic variation among languages. Therefore, the chapter devotes particular attention to these aspects of the linguistic structure of Arabic and to the way this structure is represented in the Arabic orthography. Further, in light of the importance of oral language processing skills in the acquisition of reading, the chapter also discusses Arabic diglossia: it describes the linguistic distance between Colloquial or Spoken Arabic and Standard or Literary Arabic, the primacy of Standard Arabic linguistic structures in the written form of the language, and the effect of this on several linguistic processes in literacy acquisition.

Keywords Arabic · Diacritics · Diglossia · Language · Morphology · Orthography · Phonology · Reading · Spelling · Syntax

1.1 Introduction

Arabic is the native language of approximately 300 million people worldwide and is an official language in 27 states. Also, as the language of the *Quran* it is the religious and liturgical language of all Muslims everywhere. Significantly, some local spoken variety of this language is spontaneously acquired by all native speakers as their mother tongue. This variety is known as *Spoken* (or *Colloquial*) *Arabic*, a collective term that refers to the whole range of Arabic vernaculars in numerous local dialects. These are generally classified into two regional clusters: *Eastern*

E. Saiegh-Haddad (✉)
Bar-Ilan University, Israel
e-mail: saieghe@mail.biu.ac.il

R. Henkin-Roitfarb
Ben-Gurion University of the Negev, Beersheba, Israel
e-mail: henkin@bgu.ac.il

E. Saiegh-Haddad, R. M. Joshi (eds.), *Handbook of Arabic Literacy,* Literacy Studies 9,
DOI 10.1007/978-94-017-8545-7_1,

and *Western* dialects. Eastern Arabic is spoken throughout the Fertile Crescent, in the Arabic-speaking regions of Asia, in Egypt, in the Sudan, and in partially Arabized parts of East Africa. Western Arabic is spoken in the region referred to as the *Maghreb*, including Morocco, Algeria, Tunisia, Mauritania and Libya. The regional distinction between Eastern and Western Arabic coincides with contrasting linguistic differences of phonological, morphological, phonotactic, and lexical nature, pertaining most saliently to the inflection of the imperfect verb, syllable structure, and many items of lexicon.

In contrast with the dialects, the literary varieties of Arabic, namely *Classical Arabic*, *Literary Arabic* and their modern descendant, known as (*Modern*) *Standard Arabic* (MSA), have no native speakers.[1] These literary varieties constitute the primary language of literacy,[2] namely the language children are taught to read and write at school and the only variety considered, until recently, proper for writing Arabic. As such, it is the only variety with a standardized written form. Although Spoken Arabic may be phonetically represented using the Arabic alphabet (notwithstanding some spoken sounds that have no corresponding letters) there is no consensus regarding the appropriate orthographic representation of Spoken Arabic, or even as to whether it is legitimate (culturally and ideologically) to put this non-prestigious form of the language into writing.[3]

1.2 The Structure of Arabic

1.2.1 Phonology: Consonants, Vowels, Diphthongs

The rich consonantal inventory of Modern Standard Arabic comprises 28 phonemes (two of which are actually semi-vocalic, see below). Four coronals, */s t d ð/*, represented by the letters س ت د ذ respectively, whose primary articulation involves the tongue blade and the dental-alveolar location, have phonemic counterparts characterized phonetically by a velarized co-articulation known in traditional Arabic grammar as إطْباق *ʔiṭba:q* 'covering, lidding'. Articulation of these sounds

[1] These terms have historically referred to different language varieties—*Classical Arabic* referred to the language of pre-Islamic poets; *Literary Arabic* referred to the prose language of medieval Islam, while (*Modern) Standard Arabic* refers to the modern use of this language, a descendant of the former two older forms (Bateson 2003, p. 75). The distinction, however, is not strictly adhered to.

[2] Writing in some of the colloquial prestige dialects has been noted since the fifteenth century, but most prominently since the nineteenth century in the Cairene dialect for several genres of literary prose, poetry, and drama. This 'culture of the colloquial' has been challenged and evoked some opposition and debate in Egypt (Davies 2006).

[3] Historically, Colloquial Arabic is argued by scholars to have descended from "some form of inter-tribal speech in use during the period of the [Islamic] conquests containing a greater or lesser admixture of ClA [Classical Arabic], and owe their variations to the indigenous influences" (Bateson 2003, p. 94). The popular belief that Colloquial Arabic is a direct deterioration of Classical Arabic, believed to have been the spoken language of the pre-Islamic era until spoiled by foreign substrata in the newly conquered territories, has been refuted in the light of evidence that Classical Arabic was never generally spoken (ibid).

involves raising the tongue body toward the back of the soft palate (Davis 2009, p. 636),[4] so that it "seems to fill the cavity above like a lid" (Bakalla 2007, p. 459). Additional co-articulations characterize these four phonemes, including constriction of the top of the pharynx (Al-Ani 2008, p. 599; Bakalla 2007, p. 460; Broselow 2008, p. 611; Holes 2004, p. 57). They are subsequently labeled 'pharyngealized', 'velarized', or 'emphatic', and are conventionally transcribed with a diacritic underdot */ṣ ṭ ḍ ðˤ/*. In the Arabic alphabet these phonemes are represented by the letters ص ط ض ظ respectively.[5]

These velarized emphatics share with other back consonants (velar غ */ɣ/* and خ */x/*; and uvular ق */q/*) the feature of تَفْخيم *tafxi:m* 'thickening, magnifying, emphasizing' (Bakalla 2009, p. 421) caused by the tongue raising (in the primary articulation of the latter but as a secondary co-articulation in the former). In modern dialects, all these مُسْتَعْلِيَة *mustaʕliya* 'raised' consonants (velarized and velar), also ر */r/* in many cases (Holes 2004, p. 58), tend to trigger a phonological assimilation process known as 'velarization spread' or 'emphasis spread'. This process results in the lowering and backing of neighboring vowels and in the velarization of surrounding consonants within the word, and sometimes even across a word boundary, until blocked by a high or front environment. Velarization spread may proceed forward, as in *ṣa:d* [*ṣa:ḍ*] 'to hunt', where the emphatic C_1 */ṣ/* partially assimilates the non-emphatic C_2 */d/* with respect to velarization, turning it into a *[ḍ]* allophone. Alternatively, velarization spread may proceed backward, as in *wasaṭ* [*wạṣạṭ*] 'middle', where the emphatic */ṭ/* velarizes the preceding non-emphatic */s/*, turning it into allophonic *[ṣ]*. The vowels in both cases become velarized as a result of this process. The two directions of spread have been claimed to stand in asymmetrical relation: regressive spread, like regressive assimilation in general, is more frequent and 'stronger'—it is more categorical (i.e. non-gradient) and less subject to blocking by consonants and high vowels (Davis 2009, p. 637).[6]

'Marginal' (Al-Ani 2008, p. 600) or 'secondary' emphatics, primarily */l m b/* in the vicinity of back vowels, may also trigger backing effects in many dialects. Notably, phonemic value has been claimed for secondary emphatics, such as */ṛ ṃ ḷ/* in Negev Arabic, e.g., *na:ṛ* 'fire', *ʔaṃṃ* 'mother', *xa:ḷ* 'maternal uncle', respectively. But minimal pairs cannot be established since the secondary emphatics are limited to a low vocalic environment (Davis 2009, p. 637) and are thus conditioned allophones (phonetic variants of phonemes) in contrast with the true or primary emphatic phonemes which are by definition non-conditioned. Moreover, for example, in the Negev Arabic pair *xaḷḷ-i:*(*h*)[7] 'my vinegar' vs. *xall-i:h* 'leave him' (Shawarbah 2012, p. 55), velarization in the former affects the entire lexeme [*χạḷḷ*], and a pair cannot be minimal if it differs

[4] According to other descriptions, the back of the tongue is raised towards the velum, i.e. the extreme back of the palate (Bakalla 2007, p. 459; Shawarbah 2012, p. 54).

[5] In many modern dialects, including Negev Arabic, *ḍ* and *ðˤ* have merged and are pronounced as an interdental emphatic, like the historical/*ðˤ*/.

[6] But Al-Ani (2008, p. 600) claims the opposite: "The progressive spreading is the most common, whereas regressive spreading is very rare".

[7] The 1st person sg. possessive and accusative suffixes in Negev Arabic, stressed *-i:* 'my' and *-ni:* 'me' respectively, may end in an *h*-like off-glide, so that *ʔibni:h* 'my son' is indistinguishable from the imperative *ʔibni:h* 'build it' (Blanc 1970, p. 131; Henkin 2010, p. 14).

in more than one segment. The same is true for the oft-cited 'minimal pair' *walla:h* 'he appointed him' vs. *waḷḷa:h* 'by God' (see for example, Al-Ani 2008, p. 600). Since the latter word is emphatic throughout [*wạḷḷạ:ḥ]*, the pair is far from minimal. Notably, the velarized consonant develarizes in a front environment, as in *l-illa:h* 'to God', which shows it to be a conditioned allophone. In any case, it is agreed among Arabists that the phonological scope of emphasis and rules of velarization spread are highly dialect-specific: "dialects may differ in the domain of emphasis spread, the direction of emphasis spread, the set of consonants that trigger emphasis spread, and the set of segments that block emphasis spread" (Broselow 2008, p. 610 citing Watson 2002, pp. 273–275). Moreover, the phonological scope of emphasis emanating from both 'primary emphatics', i.e. the four conventionally recognized emphatics of Classical Arabic, and 'secondary emphatics', such as /*l̩ m̩ b̩*/, is a suprasegmental phenomenon pertaining to both phonetics and phonotactics. Notably, it tends to influence the phonetic realization of consonants and vowels in MSA which, in the absence of an accepted MSA norm, will reflect the speaker's native dialect (Holes 2004, p. 58). Most importantly for our study, this spreading phenomenon results in a large set of velarized allophones. Some of these allophonic variants coincide with Arabic phonemes that have orthographic representation in the Arabic alphabet, including (ظ ط ض ص). This, as we will explain later, becomes an important issue in spelling Arabic and a source of orthographic opacity.

Two of the 28 conventional 'consonants', namely the glides /*w*/ and /*y*/, are in fact better considered semi-vowels (or semi-consonants): like consonants and unlike vowels, the glides may open a syllable (Holes 2004, p. 57); but in other respects, including the articulatory, acoustic and even orthographic (see Sect. 1.3: Orthography), they act like a prolongation of the corresponding vowels */u/* and */i/* respectively: the letter و represents both the semi-consonantal glide /*w*/ and the long vowel /*u:*/; correspondingly, the letter ي represents simultaneously the semi-consonantal glide /*y*/ *a*nd the long vowel */i:/*.

Notwithstanding the large consonantal inventory of Standard Arabic, its vocalic inventory is small, consisting of just 6 vowel phonemes. The three short vowels are low /*a*/, high front */i/*, and high back /*u*/, corresponding to their respective long equivalents: */a:/, /i:/,* and /*u:*/ (Broselow 2008, p. 609), as in *walad* 'boy', *bint* 'girl', *ʔumm* 'mother'; *na:s* 'people', *di:n* 'religion', *du:r* 'houses', respectively. In fact, some linguists (cf. Holes 2004, p. 57) recognize even fewer vocalic phonemes—just three (short) vowels, and an element of length applicable to both vowels and consonants: a geminated or lengthened consonant such as *ll* by this approach is prosodically equivalent to a long vowel, such as */a:/*. But it must be remembered that the distributional properties of lengthened vowels and geminated consonants are very different: a geminated *ll* may 'split' to two distinct, non-adjacent ones *l*V*l*. Thus, the root *DLL* gives both *dall* 'to guide' (with a geminated *ll*) and *dali:l* 'proof' (where the two root consonants C_2-*l* and C_3-*l* are separated by a vowel */i:)*. In contrast, a long vowel such as */a:/* cannot 'split' to two non-adjacent short ones, in a sequence such as *a*C*a*.

Ancient Arabic dialects, specifically eastern ones, appear to have had a fourth long vowel, the result of إمالة *ʔima:la* 'inclination, deflection', namely raising and fronting from an original /*a:*/ towards /*e:*/ or even /*i:*/ (Levin 2007; Versteegh 2001, p. 42; Wright 1975 I, p. 10). Medial (word internal) ʔ*ima:la* of several types has been

recognized in modern dialects. Minimal pairs in some sub-dialects of Negev Arabic include *jdæ:d* 'new' (plural) / *jda:d* 'forefathers', *bæ:liy* 'worn out' (participle) / *ba:li* 'my mind' (Henkin 2010, p. 53). Two secondary phonemes in many dialects are /*e:*/ and /*o:*/, resulting from diphthong contraction (see below): *mawt⇒mo:t* 'death'; *sayf⇒se:f* 'sword'.

The term 'diphthong', known in Arabic as صَوْت مُرَكَّب *ṣawt murakkab* 'compound sound', is applied in Semitic linguistics to a combination of a vowel and a glide, rather than to a sequence of two adjacent vowels forming the peak of a syllable, as in other languages. In traditional Arabic grammar just two falling diphthongs are recognized: *aw* and *ay* (al-Ani 2008, p. 599; Iványi 2006, p. 640). Widespread contraction or monophthongization of these in the dialects, especially in front phonetic environments, has given rise to two additional long vowels of Spoken Arabic, *e:* and *o:*. Both are at least partially phonemic, as witnessed by minimal pairs such as *de:r* 'monastery' vs. *di:r* 'put' (imperative); *do:r* 'turn, role' vs. *du:r* 'houses'. However, not all native speakers perceive the difference between /*e:*/ and /*i:*/, or between /*o:*/ and /*u:*/, even in dialects where some phonemic status has been established (cp. Blanc 1970, p. 118 for Negev Arabic).

1.2.2 Phonotactics: Root Structure, Syllable Structure, Stress

All 28 Arabic consonants may function as root radicals. However, there are some constraints on the distribution of some consonants, mainly on the co-occurrence of root consonants that are identical, homorganic or otherwise similar. For example, C_2 and C_3 may be identical, as in *RDD*, whence *radd* 'to return'; but C_1 and C_2 cannot be identical. A comprehensive table, devised by Greenberg (Frisch 2008, p. 625), presents the co-occurrence of all consonant groups with each other on a gradient of similarity and co-occurrence, and a principle of similarity and preference in inverse correlation. Moreover, Frisch (2008, p. 628) proposes a functional base for the principle of dissimilation, namely that similarity poses a cognitive load and is therefore undesirable: "forms without repetition are easier to produce, perceive, and hold in short-term memory". Some basic principles are as follows (Broselow 2008, p. 610):

> Generally, roots are unlikely to contain adjacent labial consonants (/*b f m*/). Adjacent coronals are avoided if they also share similar manners of articulation; thus, roots with adjacent coronal sonorants, coronal stops, or coronal fricatives are rare, and even combinations of a coronal stop and a coronal fricative are unlikely. In the posterior regions, combinations of velar and uvular consonants are avoided, as are combinations of guttural consonants.[8]

All syllables in Modern Standard Arabic begin with a single consonant (C) or glide, serving as the syllable onset and necessarily followed by a vowel (V), as the syllable nucleus or peak. The minimal syllable is thus CV, as in the preposition *li* 'to'. This is known as an open syllable, because it ends in a vowel, which is characterized by relative openness of the vocal tract. It is monomoraic, i.e. it contains one

[8] Holes (2004, p. 99) precludes homorganic non-identical root radicals in general. Exceptions include the sonorants, which can co-occur with any other consonant in any position.

mora,[9] and is thus light. Each additional mora, be it vowel length or an additional consonant, adds heaviness. A bimoraic syllable, consisting of CV: or CVC, is thus 'heavy' (Broselow 2008, p. 612; Jesry 2009, p. 388; Kager 2009, p. 344).[10] It may be open (CV: as *ma:* 'what') or closed (CVC, as *man* 'who'). Syllables with 3–4 moras, considered 'extra heavy', or 'super heavy' in this system, are limited to pausal status. One sub-class of this category is a syllable containing both a long vowel and a closing consonant (CV:C), e.g., *ba:b* 'door'—this structure may occur word-internally in special cases, such as *ʕa:m.ma* 'public' (fm.) (Holes 2004, p. 61); another is a syllable that is 'doubly' closed with two consonants: CVCC, e.g., *kalb* 'dog' or even CV:CC, e.g., *ma:rr* 'passer by'—this last type, however, is limited to geminate consonants (Broselow 2008, p. 610 ff.; Jesry 2009, p. 388).

Importantly, Arabic syllable boundaries vary with morphological processes such as declension that the words might undergo. Since syllabification in junctural (connected) prose operates across the boundaries of words in sequence, we find Standard Arabic pausal (basic) forms resyllabified in non-pausal connected or context status, e.g., pausal *jadd* 'grandfather' vs. context *jaddun* (*jad.dun*); pausal *maktab* (*mak.tab)* 'office' vs. *maktabu š-šurṭa* (*mak.ta.buš.šur.ṭa)* 'the police office' in a construct phrase. The Standard Arabic sequence *min* 'from' and *l-bayt* 'the house' potentially forms a 3-consonant cluster (*nlb*). Since Arabic does not permit 3-consonant clusters in principle, an anaptyctic (helping vowel) is inserted to break the cluster, forming *min-al-bayt* (*mi.nal.bayt)* 'from the house'.

It is noteworthy that Arabic vernaculars may vary in their syllable structure and their phonotactic constraints. For instance, Palestinian Arabic allows many 2-consonant clusters in syllable-initial positions (e.g., *tra:b* 'soil' or *kla:b* 'dogs') or across morpheme-boundaries in some grammatical forms (e.g., definite nouns *l-be:t* 'the house'). Yet, syllable final clusters are not as prevalent. The sonority principle of final anaptyxis is $C_1VC_2C_3 \Rightarrow C_1VC_2VC_3$ if Sonority $C_2 <$ Sonority C_3 (Zemánek 2006a, p. 86). In other words, a rise in sonority within a final C_2C_3 cluster will call for anaptyxis, so *qabl* 'before' (sonority rises from C_2*b* to C_3*l*) $\Rightarrow$*qabil*. Notably, the sonority hierarchy for final clusters is directly contrary to the sonority hierarchy for initial clusters, where anaptyxis is called for in the case of falling sonority. Thus, perfectly acceptable word-initial clusters of a C_1 stop or fricative and a C_2 sonorant of higher sonority, such as *dr, bl, tn, fl, sm* in ***dr****u:s* 'lessons', ***bl****a:d* 'country', ***tn****e:n* 'two', ***fl****a:n* 'so-and-so', ***sm****i:n* 'fat', will need anaptyxis in word final position, as in *ba.****dir*** 'full moon', *qa.****bil*** 'before', *ma.****tin*** 'corpus', *ṭi.****fil*** 'child', *ʔi.****sim*** 'name', respectively. Word-final clustering is more generally acceptable in the case of dropping sonority: *ʔaka****lt*** 'I/you ate', *ka****lb*** 'dog', *ħa****md*** 'praise', though again, dialects vary with respect to clustering in such cases.

Arabic stress is non-phonemic (Holes 2004, p. 62) or non-distinctive (Kager 2009, p. 344), and is predictable (though dialect-dependent), given the weight

[9] A mora is a prosodic weight unit for classifying syllable structure. It counts all units excluding the onset consonant.

[10] Holes (2004, p. 62 ff.) considers bimoraic syllables 'light' too; 'heavy' syllables in this system contain 3–4 moras. Al-Ani (2008, p. 601) similarly considers CVC a light syllable. A little further on in the article, however, Al-Ani (2008, p. 602) posits an in-between category of 'medium' or bimoraic syllables, such as *kam* 'how many' and *ma:* 'what'.

and number of syllables in the word.[11] In Standard Arabic, a word (in pausal status only) can contain just one extra-heavy syllable (of four elements or more)—that syllable is necessarily final, and receives stress, e.g., *ki.**ta:b*** 'book', *ka.**tabt*** 'I/you wrote'. In the absence of extra-heavy syllables, stress falls on the rightmost non-final heavy syllable (Kager 2009, p. 349): *mu.dar.ri.**su:.**na* 'teachers'; *yas.ta.**ʈi:.**ʕu* 'he is able'; *kas.**sar**.tu.hu* 'I broke it', ***mak**.tab* or ***mak.**ta.bun* 'office'. Otherwise, stress falls on the first syllable, e.g., ***ba.**ra.ka* 'blessing', ***ka**.ta.bu:* 'they wrote'.[12] Stress variation in Modern Standard Arabic is due, at least in part, to the fact that, as in the issue of syllable structure, here too speakers are influenced by their native dialects, which vary considerably in their stress rules. The Standard Arabic stress scheme just outlined is very similar to that of Eastern Arabic dialects (Kager 2009, p. 350).

1.2.3 Morphology: Root, Pattern[13]

Arabic, like other Semitic languages, is characterized by a predominantly non-linear or non-concatenative morphological structure (Larcher 2006; McCarthy 1981), the hallmark of which is a جَذْر *jaðr* 'root' and a derivational or inflectional pattern ميزان صَرْفيّ *mi:za:n ṣarfiyy*.

In Semitic languages, morphological derivation and inflection typically involve two bound morphemes: a triliteral (and sometimes quadriliteral) root (e.g., $_{C1}$*K*-$_{C2}$*T*-$_{C3}$*B*) and a word pattern or template (Broselow 2008, p. 610; Holes 2004, p. 99), such as $C_1a{:}C_2iC_3$ e.g., *ka:tib* 'writer' (active participle) or $maC_1C_2u{:}C_3$, e.g., *maktu:b* 'written' (passive participle). The root is an unpronounceable bound morpheme, "a skeleton of consonants" (Bentin and Frost 1995, p. 273) that provides the core meaning, or the semantic family. The pattern is a non-pronounceable bound morpheme too—a fixed prosodic template with slots for the root consonants. The insertion of the root consonants within the word pattern produces a unique lexical item with a unique meaning and a well-defined grammatical category directly discernible by the specific word pattern. It is noteworthy that while patterns are

[11] Holes ibid presents rare cases where phonemic status may be attributed to stress. This is due to neutralization of word final gemination, which results in minimal pairs such as dialectal ***sA**kat 'he was silent' vs. sa**kAt**+t⇒sa**kAt**.* 'I was/you were silent'. But he notes that such cases are "marginal and artificial".

[12] More elaborate stress rules (Holes 2004, p. 62 ff.) account for cases like *yas.**ta**.mi.ʕu* 'he listens', *muš.ki.**la**.tu.ka* 'your problem' and, particularly, when all the non-final syllables are light, e.g., *ma.li.ka.tu.hu* 'his queen'. In this case there is no general agreement as to whether the stress fell on the first syllable in Classical Arabic ***ma**.li.ka.tu.hu* (Kager 2009, p. 349), or was limited to the last three syllables (Broselow 2008, p. 613), namely *ma.li.**ka**.tu.hu*, the Arab grammarians having totally ignored the issue of stress in their writings.

[13] In the following two sections we discuss mainly Modern Standard Arabic. In demonstrating the forms, however, we choose variants that are as close as possible to those of Spoken Arabic. We thus prefer pausal forms that omit final short vowels in the same way as dialectal variants, e.g., *katab* (and not *kataba)* 'to write', Impf. *yaktub* (rather than *yaktubu),* unless the omitted vowels are the issue discussed, or when historical morpho-phonological processes are being shown, e.g., *ramaya⇒rama*: 'to throw'.

primarily vocalic templates (vowel patterns), some patterns involve gemination of root consonants or vowel length, and others are augmented with certain consonants, such as */ʔ s t n/*. In the case of verbs, these augmented patterns are called أَفْعالمَزيدَة *ʔafʕa:l mazi:da* 'augmented verbs', namely all Arabic verb patterns except for pattern I, referred to as فِعْل مُجَرَّد *fiʕl mujarrad* 'bare verb', because it consists only of the root consonants and vocalic pattern. Importantly, the additional consonants of the augmented verbs, as well as the long vowels of word patterns, are an indispensable part of the orthographic representation of words, even in unvoweled Arabic script (see Sect. 1.3: Orthography).

The root-pattern morphological structure is common to almost all Arabic content words and some function words, such as *qabl* 'before'; their semantic identity is largely determined by the consonantal root. Interestingly, even loan words, such as *talfizyo:n* 'television' and *talifo:n* 'telephone', are treated by speakers as having an internal root-pattern structure; via a derivational process known as 'root extraction', new quadriliteral roots *TLFZ* and *TLFN* are derived and combine with the quadriliteral pattern $C_1aC_2C_3aC_4$ to form the verbs *talfaz* 'to televise' and *talfan* 'to phone'. Root consonants usually preserve their phonemic identity when combining with word patterns to form Arabic lexemes. Yet, because of velarization spread (the phonological assimilation process described earlier) some root consonants may become emphatic. This phonetic change is not represented, however, in the orthographic structure of Arabic words and this may lead to orthographic opacity (see Sect. 1.3: Orthography).

All consonants, including glides, can function as root-radicals. A root containing a glide, however, is considered مُعْتَلّ *muʕtall* 'weak',[14] being prone to morpho-phonological changes. These contrast with the 'strong' or 'sound' roots called صَحيح *ṣaħi:ħ* 'correct' whose radicals remain phonologically stable (Akesson 2009, p. 121; Holes 2004, p. 110 ff.; Versteegh 2001, p. 85 ff.; Versteegh 2007b, p. 309). In a C_1-glide root, known as مِثال *miθa:l* 'assimilated', e.g., *WJD* 'find', the glide may be elided in the Impf. **yawjidu* ⇒*yajidu* 'he finds'; a C_2-glide root, known as أَجْوَف *ʔajwaf* 'hollow', e.g., *QWL*, undergoes several changes, e.g., **qawaltu* ⇒*qultu* 'I said', Impf. **ʔaqwulu* ⇒*ʔaquwlu* ⇒*ʔaqu:lu* 'I say'; **qawalat* ⇒*qa:lat* 'she said'; a C_3-glide root, known as ناقِص *na:qiṣ* 'defective', such as *RMY,* is also prone to morpho-phonological changes, e.g., **ramaya* ⇒*rama:* 'to throw', Impf. **yarmiyu* ⇒*yarmiy* ⇒*yarmi:* 'he throws' (Akesson 2009, pp. 121–122; Chekayri 2007, p. 164 ff.).[15]

Most traditional Arabic dictionaries are alphabetically ordered by consonantal roots and they specify in each entry the specific meaning that results from the

[14] Some scholars include hamzated verbs, i.e. verbs containing *hamza* (see Sect. 1.3: Orthography), in the category of weak verbs (e.g., Voigt 2009, p. 700 ff.).

[15] The grammarians set up phonotactic rules according to a scale of relative lightness and strength of the phonemes that corresponds to sonority (Holes 2004, p. 113): vowels are lightest and strongest, consonants heaviest and weakest; within the vowels, the hierarchy is *a*>*i*>*u*. In contact, the lighter-stronger phoneme overrules and only sequences of rising lightness are permitted. So the triphthong *iyu* in **yarmiyu* above will contract to *iy* ⇒*i:*, as also in **qa:ḍiyu* ⇒*qa:ḍi:* 'judge' (Versteegh 2001, p. 86 ff.; Voigt 2009, p. 699). The homogeneous triphthongs **awa,***aya* are simplified by elision of the glide, as we saw in **qawala* ⇒*qa:la* and **ramaya* ⇒*rama*: above.