Applied Psycholinguistics **25** (2004), 495–512
Printed in the United States of America
DOI: 10.1017.S0142716404001249

The impact of phonemic and lexical distance on the phonological analysis of words and pseudowords in a diglossic context

ELINOR SAIEGH–HADDAD
Bar-Ilan University

ADDRESS FOR CORRESPONDENCE
Dr. Elinor Saiegh–Haddad, English Department, Bar-Ilan University, Ramat-Gan 52900, Israel.
E-mail: saieghe@mail.biu.ac.il

ABSTRACT
The study examined the impact of the phonemic and lexical distance between Modern Standard Arabic (MSA) and a spoken Arabic vernacular (SAV) on phonological analysis among kindergarten ($N = 24$) and first grade ($N = 42$) native Arabic-speaking children. We tested the effect of the lexical status of the word (SAV, MSA, and pseudoword), as well as the linguistic affiliation of the target phoneme (SAV vs. MSA), on initial and final phoneme isolation. Results showed that, when words were composed of SAV phonemes only, the lexical status of the word did not affect phoneme isolation. However, when MSA and pseudowords encoded both SAV and MSA phonemes, kindergarteners found MSA words significantly more difficult to analyze. Comparing children's ability to isolate SAV versus MSA phonemes revealed that all children found MSA phonemes significantly more difficult to isolate. Kindergarteners found MSA phonemes that were embedded within MSA words even more difficult to isolate. Results underscore the role of the lexical status of the stimulus word, as well as the linguistic affiliation of the target phoneme in phonological analysis in a diglossic context.

Characterizing the construct of phonological awareness and explicating its role in the acquisition of reading skill have been the subject of extensive theoretical and empirical attention for the last couple of decades. The evidence that this pursuit has produced establishes a strong empirical basis for the role of children's phonological awareness in the acquisition of basic reading processes in an alphabetic orthography (for reviews, see Adams, 1990; Goswami & Bryant, 1990; Wagner & Torgesen, 1987). As such, reading is grafted onto oral language skills. Therefore, a basic prerequisite to word decoding is to become aware of the phonological structure of words.

Models of reading that highlight the role of oral language skills in the acquisition of basic reading processes (Gough, Hoover, & Peterson, 1996; Gough & Tunmer, 1986; Hoover & Gough, 1990) are grounded in the premise that, by the time children begin to read in their first language (L1), they have already acquired that language in its oral mode. A second premise, and one that follows from the former,

is that all children are familiar, and to a comparable degree, with the phonological structures of their native language, the language they will learn to decode. These two basic assumptions are not always valid, however. This is at least not so in learning to read in a second language (L2) or in reading in a diglossic context. Learning to read in these contexts is not typically predicated on fully mastered oral language skills, nor on comparable familiarity among beginning readers with the various phonological structures of the language. Individual differences in the amount of exposure to, and familiarity with, oral language structures are expected, therefore, to result in individual variations in reading skill development in these populations. This hypothesis has only recently begun to attract attention, and it was first addressed in the acquisition of basic reading and spelling processes in the L2 of bilingual children. For instance, Wade–Woolley & Geva (2000) tested sensitivity to a phonemic contrast (/*ts*/ vs. /*s*/), which occurs productively in Hebrew but is phonotactically constrained in English, among English-speaking children learning Hebrew as an L2. The results showed that novel phonological elements that are specific to the L2 present additional challenges to beginning readers. Wang and Geva (2003) compared the spelling development of two novel English phonemes (/ʃ/ and /θ/) among Cantonese-speaking English as an L2 children and English as a first language children. They found that the spelling errors of the Chinese English as an L2 children reflected difficulty in representing phonemes that are absent in Cantonese phonology. Crosslinguistic differences in characteristics of the oral language were also used to explain crosslinguistic variations in the rate and the route of basic reading skill development in children from different native language backgrounds (e.g., Caravolas & Bruck, 1993; Cossu, Shankweiler, Liberman, Katz, & Tola, 1988; Durgunoglu & Oney, 1999; Oney & Goldman, 1984).

The impact on reading acquisition of children having to come to terms with novel linguistic structures that are not within their oral language experience has also been addressed in speakers of African American Vernacular English learning to read Standard English, or School English. It was reasoned that, given the phonological and grammatical disparity between the two forms, greater familiarity with the latter might be associated with more successful reading acquisition (Labov, 1995; Rickford & Rickford, 1995; Washington & Craig, 2001). This hypothesis recently acquired direct empirical evidence (Charity, Scarborough, & Griffin, 2002). Thus, research into the acquisition of reading in an L2 and in a standard dialect converge on demonstrating that linguistic disparity between the oral language of children and the written language does not support the natural development of reading.

Another issue that has concerned research in reading skills development has been the nature of the phonological awareness construct. However, research in this domain has primarily focused on the role of the phonological structure of the stimulus word and of the target phonological unit on phonological analysis (e.g., word length, linguistic context, size of phonological unit, articulatory/acoustic features of target phoneme, etc.; for a review, see McBride–Chang, 1995). The impact on phonological analysis of the lexical status of the stimulus word (word vs. nonword or pseudoword), and of aspects of the lexical representation of the word (such as word frequency, word familiarity, neighborhood density, etc.) has only recently begun to attract attention. The lexical restructuring model (Metsala &

Walley, 1998; Walley, Metsala, & Garlock, 2003) was developed as an attempt to capture the relevance of various aspects of the lexical representation of the word to phonemic awareness and early reading. According to this model, spoken vocabulary growth affects the quality of the phonological representation of words, because it increases the pressure for more fine-grained segmental representation (Elbro, 1996; Fowler, 1991; Metsala, 1997a, 1997b, 1999; Walley, 1993; Walley & Fledge, 1999). As a result, familiar words should have better segmented representations much earlier than unfamiliar or pseudowords.

DIGLOSSIA AND THE ACQUISITION OF BASIC READING PROCESSES IN ARABIC

The term *diglossia*, which was first introduced by Ferguson (1959), describes a situation that includes the following features: (a) a differentiation between the written and the oral modes; (b) a rigid sociofunctional complementarity of two separate sets of functions performed by two different linguistic codes; (c) a rich and dominant written literary tradition; and (d) linguistic relatedness between the two linguistic codes: the written and the spoken. Arabic is a clear case of diglossia. The Arabic language (*al'arabiyya*) refers to a hierarchy of various language registers, including modern standard Arabic (MSA) and all spoken Arabic vernaculars (SAVs) used in the Arabic-speaking world; it is "a symbolic abstraction comprising old and new language norms and standards of all the linguistic varieties of Arabic" (Maamouri, 1998, p. 30). Spoken Arabic vernaculars are all linguistically related to MSA. Nonetheless, they are remarkably distinct from it. Though the nature and the linguistic manifestation of the distance between the two forms may be different in different vernaculars, a comparison between MSA and any given SAV always reveals phonological, morphosyntactic, and lexical–semantic differences.

A major manifestation of the linguistic distance between MSA and a given SAV pertains to the phonemic inventory. Although MSA shares most of the phonemes with all spoken vernaculars, no single SAV has the same set of phonemes as MSA (Maamouri, 1998). As a consequence, a specific word might have a phonological representation in SAV that, although phonologically related to, differs from its phonological representation in MSA (e.g., MSA /θaʕlab/, meaning "a fox," versus SAV /ṯaʕlab/). At the level of the lexicon, MSA and SAV share many words. These words, however, exhibit variable degrees of phonological disparity between their MSA and SAV forms. This disparity may be minute, as the deletion of MSA inflectional morphemes, MSA /kaṯaba/ ("he wrote") versus SAV /kaṯab/, or quite remarkable, as in altering the segmental structure of the word, as in MSA /ðanab/ ("a tail") versus SAV /danab/; the prosodic (syllabic) structure of the word, as in MSA /sahl/ ("a plain") versus SAV /sahel/; or both the segmental and the prosodic structure of the word, as in MSA /θalʒ/ ("snow") versus SAV /ṯaleʒ/. Only few words have identical phonological representations in both MSA and SAV. In addition to shared lexical items like those just illustrated, the majority of words, including function words and many high-frequency content words, have unique lexical representations in MSA and are, thus, to be learned in the process of becoming literate in the standard form of the language.

The phonological distance, yet relatedness, between the forms of words in MSA and SAV results in another peculiar feature of the lexical representation of words that is unique to a diglossic context. In Arabic, MSA words may be composed of SAV phonemes only, in which case lexical affiliation to MSA is based on the lexical representation of the word only. For example, the word /ħaṭ/ (he put) is the SAV form of the MSA verb /wadaʔa/, even though both are composed of SAV phonemes only. Alternatively, a given MSA word may encode MSA phonemes as well, in which case, even though a phonologically related form of the word may exist in SAV, lexical affiliation to MSA is justified by the presence of at least a single MSA phoneme. For instance, the word "gold" in Arabic has two forms: the MSA form /ðahab/ and the SAV form /dahab/, which differ only in the linguistic affiliation of the initial phoneme: MSA /ð/ versus SAV /d/. This unique characteristic, in which lexical affiliation is based not only on the lexical representation of the word but also on its segmental structure, allows a linguistic unpacking of lexical familiarity, or word knowledge in Arabic. In other words, unfamiliarity with a specific lexical item can be traced back to at least three sources: (a) lack of familiarity with the lexical representation of the word, which usually coincides with lack of familiarity with the phonological representation of the word as well; (b) lack of familiarity with one or more of the component phonemes that make up the word; or (c) lack of familiarity with both the lexical as well as the segmental structure of the word. This context provides, thus, an ideal environment for testing the simultaneous effect of the phonemic and lexical aspects of the word on phonological analysis.

Despite the peculiar nature of Arabic diglossia and the unique natural setting that it furnishes for an exploration of the psychology and linguistics of literacy acquisition, work into this area has been phenomenally restricted. The very few pertinent studies tested the role of exposure to MSA on the development of literary language skills and reading comprehension among children. They found that exposure to MSA was significantly correlated with the acquisition of reading comprehension in MSA (Abu-Rabia, 2000; Feitelson, Goldstein, Iraqi, & Share, 1993). Given the documented positive impact of oral exposure to MSA on children's reading comprehension, it was hypothesized that familiarity, or lack of it, with specific and explicitly defined MSA linguistic structures would affect the acquisition of basic bottom-level processes that underlie reading comprehension. Pursuing this question was expected to make two important contributions to reading research in a diglossic context. The first is to force researchers to become explicit about bottom-level diglossic structures and to test their direct effect on the acquisition of basic reading processes. The second is to encourage researchers to address specific reading processes and to probe their sensitivity to various diglossic linguistic structures. This twofold objective has been pursued in a programmatic project, the present study is part of which, that tested the effect of diglossia on the acquisition of basic reading processes in Arabic. Saiegh–Haddad (2003) examined the effect of the phonological distance between MSA and a local form of Palestinian SAV spoken in the north of Israel on the acquisition of phonemic awareness and word decoding skills among kindergarten and first-grade children. Two phonological structures were addressed: the phoneme and the word syllabic structure. The results showed that both kindergarten and first-grade children had

particular difficulty isolating MSA phonemes. Further, children found it more difficult to isolate a MSA phoneme that was embedded within an MSA syllabic structure. Also, despite high reading accuracy rates, MSA phonemes and MSA syllabic structures were found to be a major source of decoding inaccuracy among children.

The present study is part of the same programmatic research project that aims to explore the role of the linguistic disparity between MSA and SAV on the acquisition of basic reading processes in diglossic Arabic. Previous research has demonstrated that MSA phonological structures are more difficult for children to analyze and to decode than SAV structures. This evidence, however, comes from the analysis of pseudowords only (Saiegh–Haddad, 2003). As a result, it is still unclear whether the lexical status of the word affects phonological analysis as well. Further, it is not clear whether lexical status has the same impact on the analysis of phonemes with different linguistic affiliation: SAV versus MSA phonemes. The present study addresses these questions. It asks whether the lexical distance between MSA and SAV affects the phonological analysis of children. Further, it examines the simultaneous effect of both the lexical status of the word, as well as the linguistic affiliation of the target phoneme on phonological analysis.

Three factors helped generate the next hypothesis. The first is the substantial role that current reading theory attributes to oral language skills in the acquisition of basic reading processes (Adams, 1990; Goswami & Bryant, 1990; Gough & Tunmer, 1986; Hoover & Gough, 1990). The second is the unique reading problems observed among bilingual readers (Wade–Woolley & Geva, 2000; Wang & Geva, 2003) and among readers of a standard dialect (Charity et al., 2002), which are arguably triggered by linguistic disparity between the oral language of children and the written language. The third is the documented positive effect of word familiarity on segmental analysis among children (Metsala & Walley, 1998; Walley et al., 2003). Hence, it is hypothesized that novel MSA linguistic structures that are not within children's SAV may not support the natural development of basic reading processes in MSA. Two manifestations of the linguistic distance between MSA and SAV are tested: phonemic and lexical. The following questions are addressed:

1. Does a word's lexical status (SAV, MSA, or pseudoword) affect children's SAV phoneme isolation performance?
2. What is the effect of a word's lexical status (MSA vs. pseudoword) on the isolation of phonemes with different linguistic affiliation (SAV phonemes vs. MSA phonemes)?

These questions were tested with kindergarten and first-grade children. It was predicted that SAV structures, both phonemes and words, would be easier to analyze than MSA structures. In turn, both MSA and SAV words should be easier to analyze than pseudowords. Further, because formal and consistent exposure to MSA begins in the first grade, first-grade children were expected to show an advantage over kindergarteners when novel MSA structures (words and phonemes) are analyzed. Finally, it was predicted that, in line with previous evidence from

Arabic (Saiegh–Haddad, 2003), final phoneme isolation would be easier than initial phoneme isolation. However, the position of the target phoneme was not expected to interact with the effect of the lexical status of the word, or the linguistic affiliation of the phoneme.

METHOD

Participants

A total of 66 children (24 kindergarten, 42 first grade) from a local public school in a small Arab village in the north of Israel participated in the study. Children were all middle class Arabs and native speakers of the same SAV. Those speaking a different SAV, because they were new arrivals to the village, were not tested. The mean age of children in the kindergarten group was 5;9 ($SD = 3.30$ months), while the mean age of the first-grade group was 6;10 ($SD = 3.77$ months).

Materials

Children were given two phonemic awareness test clusters: initial phoneme isolation (50 items) and final phoneme isolation (50 items). Each cluster was made up of three tests that differed in the lexical status of the stimulus word. SAV words were high frequency words that had identical phonological representations in both SAV and MSA. MSA words were drawn from the first grade primer and combined both MSA words that had phonologically related forms in children's SAV and words that had a unique lexical representation in MSA. None of the MSA words had an equivalent in the children's SAV. Concerning the pseudowords, it is important to point out that SAV words can never encode MSA phonemes. Otherwise, they would be categorized as MSA words. On the other hand, MSA words and pseudowords may or may not encode MSA phonemes. As a consequence, the study employed two classes of MSA and pseudoword stimuli, words composed of SAV phonemes only and words that encoded one MSA phoneme, in either a word-initial or word-final position. All words were disyllabic (consonant–vowel–consonant–vowel–consonant [CVCVC]) and three to four letters long (see Appendixes A and B). The study targeted all the MSA phonemes that were not part of the phonemic inventory of the participants' SAV. These phonemes were evenly distributed between initial and final positions and included both the voiced and the voiceless interdental fricative phonemes, /ð/ and /θ/, like the first phonemes in the English words "then" and "thin," respectively; the emphatic voiced interdental fricative /$\underline{ð}$/; and the uvular stop /ɑ/. SAV phonemes included a random sample of all the consonantal phonemes that existed in children's SAV and included both obstruents (stops and fricatives) and sonorant consonants (nasals and liquids).

Procedure

The administration of all tasks took place during the last month of the school year. The order of the administration of initial and final phoneme isolation tasks was

Table 1. *SAV phoneme isolation by word's lexical status, phoneme position, and grade: Summary statistics*

		Kindergarten ($N = 24$)		First Grade ($N = 42$)		Both Grades ($N = 66$)	
Word's Lexical Status		Initial	Final	Initial	Final	Initial	Final
SAV	*M*	6.26	8.20	8.33	9.73	7.59	9.18
(Max. 10)	*SD*	4.02	2.53	3.07	0.66	3.52	1.76
MSA	*M*	6.00	8.12	8.21	9.64	7.40	9.09
(Max. 10)	*SD*	4.00	2.98	2.92	0.75	3.49	2.01
Pseudowords	*M*	6.33	8.00	8.02	9.45	7.40	8.92
(Max.10)	*SD*	3.44	2.79	3.23	0.77	3.38	1.90
Total	*M*	18.62	24.33	24.57	28.83	22.40	27.19
(Max. 30)	*SD*	9.89	8.03	8.70	1.44	9.52	5.37

Note: SAV, spoken Arabic vernacular; MSA, modern standard Arabic.

counterbalanced. SAV words were administered first, followed by MSA words and then pseudowords. MSA words and pseudowords encoding and not encoding MSA phonemes were randomly ordered. The same child was administered all phonemic awareness tasks on the same day. Children were asked to listen carefully to the stimuli that were presented separately using an audiotape. They were asked to repeat each stimulus word out loud, and only then to vocalize the target phoneme. This procedure was followed in order to prevent pronunciation problems from becoming confounded with an inability to isolate the target phoneme. Children who were able to pronounce MSA phonemes properly participated in the study. Two illustration trials using SAV words and SAV target phonemes were administered to familiarize children with the task. The participants' responses were recorded and also noted on a sheet of paper that was only used by the examiner for cross-validation.

Each participant received six overall phonemic awareness scores that corresponded to the number of correct responses achieved on each of the tests. In addition, four subscores were produced per each position category (initial vs. final) that corresponded to the number of correct responses that participants achieved on SAV and MSA phonemes separately. One point was assigned for successfully isolating the target phoneme from the stimulus word, and a zero score was assigned for providing a wrong phoneme or phonemes or for providing the correct phoneme embedded within a CV syllabic unit.

RESULTS

Table 1 provides summary statistics of children's SAV phoneme isolation performance by word's lexical status (SAV, MSA, and pseudowords), phoneme position (initial and final), and grade (kindergarten and first grade).

Table 2 and Table 3 provide descriptive statistics of children's phoneme isolation as a function of word's lexical status (MSA vs. pseudoword), phoneme's linguistic

Table 2. *Kindergarteners' (N = 24) phoneme isolation by word's lexical status, phoneme's linguistic affiliation, and phoneme position: Summary statistics*

Phoneme's Affiliation	Phoneme Position								
	Initial			Final			Total		
	SAV	MSA	Both	SAV	MSA	Both	SAV	MSA	Both
Maximum Score Word Status	(10)	(10)	(20)	(10)	(10)	(20)	(20)	(20)	(40)
MSA									
M	6.00	3.45	9.45	8.12	7.00	15.12	14.12	10.45	24.58
SD	4.00	3.23	6.80	2.98	3.23	5.58	4.23	4.73	8.10
Pseudowords									
M	6.33	5.04	11.37	8.00	6.37	14.37	14.33	11.41	33.57
SD	3.44	3.11	5.98	2.79	2.76	4.65	4.01	4.96	6.90
Total									
M	12.33	8.50	20.83	16.12	13.37	29.50	28.45	21.86	58.15
SD	6.42	5.67	11.30	5.59	5.71	9.94	7.19	9.22	14.37

Note: SAV, spoken Arabic vernacular; MSA, modern standard Arabic.

Table 3. *First graders' (N = 42) phoneme isolation by word's lexical status, phoneme's linguistic affiliation, phoneme position, and grade: Summary statistics*

Phoneme's Affiliation	Phoneme Position								
	Initial			Final			Total		
	SAV	MSA	Both	SAV	MSA	Both	SAV	MSA	Both
Maximum Score Word Status	(10)	(10)	(20)	(10)	(10)	(20)	(20)	(20)	(40)
MSA									
M	8.21	7.47	15.69	9.64	9.71	19.35	17.85	17.19	35.04
SD	2.92	3.27	5.94	0.75	1.79	2.20	3.11	4.72	7.37
Pseudowords									
M	8.02	7.26	15.28	9.45	8.83	18.28	17.47	16.09	33.57
SD	3.23	3.02	6.08	0.77	1.36	1.70	3.47	3.81	6.90
Total									
M	16.23	14.73	30.97	19.09	18.54	37.64	35.33	33.28	68.61
SD	6.01	6.16	11.86	1.26	2.81	3.51	6.38	8.35	14.09

Note: SAV, spoken Arabic vernacular; MSA, modern standard Arabic.

affiliation (SAV vs. MSA phoneme), and phoneme position (initial vs. final) for kindergarteners and first graders, respectively.

To test the role of a word's lexical status on the isolation of phonemes with different linguistic affiliation: SAV and MSA phonemes, two independent one-way analyses of variance (ANOVAs) were conducted. The first addressed children's isolation of SAV phonemes and analyzed their performance on SAV, MSA, and pseudowords that encoded SAV phonemes only. The second addressed children's isolation of SAV versus MSA phonemes and analyzed their performance on MSA and pseudowords only. SAV words were not included in the latter analysis because these words can be composed of SAV phonemes only.

A $3 \times 2 \times 2$ ANOVA with repeated measures was used to test the effect of word's lexical status (SAV, MSA, and pseudoword) and phoneme position (initial vs. final) on SAV phoneme isolation. Grade (kindergarten vs. first grade) was also included in the analysis as a between-subjects variable. The main effect of a word's lexical status on the isolation of SAV phonemes was not significant. However, the main effects of grade as a between-subjects factor and of phoneme position as a within-subjects variable were significant: grade, $F\ (1, 63) = 17.56$, $p < .001$; phoneme position, $F\ (1, 63) = 11.94$, $p < .001$; first grade children achieved significantly higher phoneme isolation scores than kindergarteners and initial phonemes were significantly more difficult to isolate than final phonemes. The interaction of word's lexical status by grade was not significant.

To address the impact of a word's lexical status on the isolation of phonemes with different types of linguistic affiliation, children's phonological analysis of MSA and pseudowords was analyzed using a $2 \times 2 \times 2 \times 2$ ANOVA with repeated measures on word's lexical status (MSA vs. pseudowords), phoneme's linguistic affiliation (SAV vs. MSA), and phoneme position (initial vs. final) and with grade (kindergarten vs. first grade) as a between-subjects factor. Again, the main effect of a word's lexical status was not significant. However, the main effect of phoneme's linguistic affiliation was significant, $F\ (1, 64) = 29.62$, $p < .001$, and MSA phonemes were significantly more difficult to isolate than SAV phonemes. The main effects of grade and of phoneme position were again significant: grade, $F\ (1, 64) = 25.32$, $p < .001$; phoneme position, $F\ (1, 64) = 22.80$, $p < .001$; first graders outperformed kindergarteners and initial phonemes were significantly more difficult to isolate than final phonemes.

The analysis also revealed a significant two-way interaction of word's lexical status by grade, $F\ (1, 64) = 6.35$, $p < .01$ (kindergarteners had particular difficulty analyzing MSA words); word's lexical status by phoneme position, $F\ (1, 64) = 7.21$, $p < .01$ (the initial phonemes of MSA words were significantly more difficult to isolate than the final phonemes); phoneme's linguistic affiliation by grade, $F\ (1, 64) = 8.18$, $p < .01$ (kindergarteners found it significantly more difficult to isolate MSA phonemes); and phoneme's linguistic affiliation by phoneme position, $F\ (1, 64) = 4.32$, $p < .05$ (MSA phonemes were significantly more difficult to isolate when embedded in an initial word position).

Finally, a three-way interaction of word's lexical status by phoneme's linguistic affiliation by phoneme position was significant, $F\ (1, 64) = 12.92$, $p < .001$, and MSA phonemes were significantly more difficult to isolate if they formed the onset of MSA words. The interaction of word's lexical status by phoneme's linguistic

affiliation by grade was nearly significant, $F\ (1, 64) = 3.21$, $p = .07$, revealing a tendency for kindergarteners to show particular difficulty when faced with the task of isolating MSA phonemes embedded within MSA words.

DISCUSSION

The first research question addressed the effect of the lexical status of the word on SAV phoneme isolation. The results of the study showed that when SAV phonemes were targeted, the lexical status of the stimulus word (SAV, MSA, or pseudoword) did not affect phoneme isolation performance. These findings provide apparent counter evidence to the predictions of the lexical restructuring model (Metsala, 1999; Metsala & Walley, 1998; Walley, 1993; Walley et al., 2003), according to which the phonological analysis of words should be better than that of pseudowords and the phonological analysis of familiar words should be better than that of unfamiliar words. Two unique features of Arabic are proposed to explain this unexpected finding, the phonological/prosodic structure of Arabic and Arabic orthography. As to the phonological/prosodic structure, it is suggested that the rather simple syllabic structure of Arabic may account for the absence of a word familiarity effect on phoneme segmentation. As such, the basic and most frequent syllable structure in Arabic is the CV, followed by the CVC structure. Consonantal cluster rime coda structures (CVCC) are very rare in MSA (Ababneh, 2000), as they are restricted to monosyllabic words and to short vowels. Except for the CVCC structure, no complex onset or coda subsyllabic units exist in MSA. Contrary to Arabic, the prosodic structure of English features a variety of complex onset and coda subsyllabic units that may embody as many as three consonants, as in the word "strands" /*strandz*/, and even a marginal length of four consonantal phonemes, including the appendix, as in "sixths" /*siksθs*/. This is believed to make the segmental analysis of English words harder, more variable, and perhaps, as a consequence, more strongly dependent on lexical aspects of the word than is the case in Arabic. Thus, it is possible that the more complex the prosodic structure of the language is, the more difficult the phoneme segmentation task will be and the more dependent it will be on the lexical status of the stimulus word. In order to probe the validity of this argument, research into the role of lexical involvement in segmental analysis should take a linguistic perspective on the prosodic structure of the stimuli used. It is possible that lexical involvement would figure more prominently in the phonemic analysis of complex prosodic structures. As such structures are rare in Arabic, the present study tested the segmental analysis of CVCVC words only. It was found that the lexical status of the word did not affect phoneme isolation. This finding, however, does not automatically invalidate the relevance of the lexical status of the word to phonological analysis in other languages, or even in Arabic but among younger children. The role of the lexical status of the word may interact with its prosodic structure in affecting phoneme segmentation. This issue remains for future research to pursue, and it would be best tested in a language that features a variety of prosodic structures and with varying degrees of complexity.

Work in English has shown that even first graders' phoneme segmentation benefited from familiarity with the stimulus word (Metsala, 1999). The present study demonstrated the absence of a word familiarity effect on the segmental

isolation of SAV phonemes even among kindergarteners. A second and tentative interpretation for this finding in Arabic relates to the nature of vowelized Arabic orthography. Not only is the prosodic structure of MSA simple, but the orthography that beginning readers are exposed to is highly transparent: shallow vowelized Arabic. Upon the inception of reading, children receive positive and reliable feedback from the orthography as to word pronunciation. This results in greater confidence in the orthography and a sharper awareness of the phoneme, as decoding is predicated primarily on a grapheme–phoneme conversion mechanism. In line with this argument, crosslinguistic research that compared the development of phonemic awareness in beginning readers of a shallow orthography, like Turkish and German, with that of beginning readers of English showed that exposure to a shallow orthography accelerates the acquisition of phonemic awareness (Caravalos & Bruck, 1993; Cossu et al., 1988; Durgunoglu & Oney, 1999; Mann & Wimmer, 2002; Oney & Goldman, 1984). Such enhanced phoneme awareness may explain the absence of lexical involvement in phoneme isolation in Arabic. The extent to which children from different oral language background and different orthographies might make use of lexical processes in phonological analysis, and the point in the acquisition process when this ceases to constitute a factor is an issue for future crosslinguistic research to explore.

The second research question addressed the effect of a word's lexical status on the isolation of phonemes with different linguistic affiliation: SAV versus MSA phonemes. The results of the study showed that the linguistic affiliation of the target phoneme, SAV versus MSA, had a main effect on phoneme isolation; MSA phonemes, that were absent in the oral language experience of children, were more difficult for both kindergarten and first grade children to isolate than SAV phonemes. At the same time, we noted a developmental effect, as MSA phonemes proved significantly more challenging for kindergarteners than for first grade children. As to the effect of a word lexical status on phoneme isolation, the results again failed to support a main effect of word familiarity on phoneme isolation: children showed comparable levels of phoneme isolation performance for MSA and pseudowords. The results also showed that, contrary to expectations, kindergarten children had more difficulty with the analysis of MSA words than pseudowords and with the isolation of MSA phonemes embedded within MSA than within pseudowords. Such a reversed effect of word familiarity than would be predicted from theories of lexical involvement in phonological analysis is argued to result from another unique linguistic feature of a diglossic context, and it sheds light on the linguistic components that go into word familiarity in Arabic. This feature is the phonological distance, yet relatedness between the phonological representations of the same word in SAV and MSA. As such, first graders, as they have had a whole year of consistent daily exposure to MSA, have internalized a rather robust phonological representation for MSA phonemes and MSA words, especially as the latter had been drawn from the first grade primer. Such high quality phonological representation (Elbro, 1996, 1998) of MSA phonemes explains the developmental effect reflected in the significantly higher MSA phoneme isolation scores among first graders. Such robust phonological representation is probably also responsible for the absence of a word familiarity effect on phoneme isolation among first graders. Contrary to first graders, the phoneme isolation performance of kindergarteners was found sensitive to both the linguistic

affiliation of the target phoneme and the lexical status of the stimulus word. The results showed that kindergarteners had particular difficulty isolating MSA phonemes. Further, while they were shown to have particular difficulty with the analysis of MSA words in general, they found MSA phonemes that were embedded within MSA words more difficult to isolate than those embedded within pseudowords. Children at the kindergarten level, and in the absence of consistent formal exposure to MSA, have probably acquired a rather vague, inaccurate, and unstable phonological representation of MSA phonemes and MSA words, the kind that is implicitly gained from occasional exposure to TV programs that use MSA, or from story reading either at school or at home. As a result, when they were required to analyze an MSA word, they had automatically activated the word's SAV phonological representation. As the two forms of the word are phonologically related, such phonological relatedness interfered with phonological analysis and resulted in lower phoneme isolation scores. For instance, when children were required to isolate the initial MSA phoneme /θ/ from the MSA word /θaʕ/ab/ ("a fox"), they automatically activated the SAV representation of this word /ʈaʕ/ab/. This resulted in a phoneme isolation error. In the absence of a related competing phonological representation in the mental lexicon of children, as in the case of pseudowords, kindergarten children achieved higher phoneme isolation scores.

Thus, while first graders may have used their rather robust knowledge of the phonological representation of MSA phonemes and MSA words to facilitate MSA phoneme isolation, kindergarteners showed a reversed pattern, with MSA words posing a particular stumbling block. The absence of a robust phonological representation of MSA words, especially given a competing overlapping phonological representation for some MSA words, disrupted the phonetic recoding of the word in working memory (Wagner & Torgesen, 1987) and led to a phoneme isolation error. In the analysis of pseudowords, and in the absence of a competing phonological representation, children achieved higher phoneme isolation scores. This finding has important implications for phonological analysis in bilingual contexts as well, especially when the two languages of bilingual children are typologically related and, thus, share a large set of phonologically overlapping cognates. The results of the study suggest that in the phonological analysis of cognates both forms of the cognate may be activated. This may interfere with the outcome of the phonological analysis process. In order to tap into the processes that underlie the outcome of phonological analysis, future research should employ introspective methods of data collection. Only such an approach would allow an examination of the factors that underlie phonological analysis and outcome. Such an approach would also inform about the mechanisms that underlie the analysis of phonologically related linguistic forms, especially as transfer may surface more prominently when the reader can make "equivalence classifications" (Fledge, 1987) on the two phonological forms available to him.

The present findings demonstrate the centrality of two manifestations of the linguistic distance between MSA and SAV in understanding the phonological sensitivity of children in a diglossic context toward the language they learn to decode. The first is the phonemic distance between MSA and SAV. The second is the lexical distance between the two forms. The results of the present study

replicate previously reported findings of the role of phonemic distance in the acquisition of phonemic awareness in a diglossic context (Saiegh–Haddad, 2003). However, rather than testing the phonological analysis of pseudowords only, the present study tested the effect of phonemic distance on the phonological analysis of different types of words: SAV, MSA, and pseudowords. Here too, the linguistic affiliation of the phoneme had a main effect on phoneme isolation, hence strengthening previously reported evidence and consolidating the contention that the availability of the phoneme in the oral language of children facilitates phoneme isolation, irrespective of the lexical status of the stimulus word within which the phoneme is embedded. These findings demonstrate the unique complexities that inhere in the acquisition of reading in a phonologically distant language (Charity et al. 2002; Saiegh–Haddad, 2003, 2004; Wade–Woolley & Geva, 2000; Wang & Geva, 2003). Children acquiring reading in a language that is phonologically distant from their oral language, as in reading in a bilingual context, in a standard dialect, and in a diglossic context, are required to construct mental representations for a set of novel phonological structures. Until such novel structures are acquired, and until their phonological representation becomes robust, and their activation automatic (LaBerge & Samuels, 1974; Perfetti, 1985; Wolf & Katzir–Cohen, 2001), novel phonological structures may well constitute an additional challenge for beginning readers.

It is noteworthy that the study tested the isolation of phonemes from both an initial and a final word position. The results showed that the isolation of initial onset phonemes was consistently more difficult than the isolation of final phonemes. This pattern was evident among both kindergarten and first-grade children, across both SAV and MSA phonemes, and across all types of words: SAV, MSA, and pseudowords. Further, initial phoneme isolation was found more sensitive than final phoneme isolation to variations in the linguistic affiliation of the target phoneme (SAV vs. MSA) and in the lexical status of the word (SAV, MSA, and pseudo). These findings replicate and extend previously reported findings from the phonemic awareness of Arabic-speaking children (Saiegh–Haddad, 2003) and show that Arab children find it more difficult to isolate the prevocalic consonantal phoneme than a postvocalic final consonant. Saiegh–Haddad (2003) found that final phonemes were easier to isolate than initial phonemes even when they were embedded in a complex consonantal cluster rime coda. The particular difficulty that Arab children revealed in isolating initial-onset phonemes than final rime coda is at odds with the patterns thus far revealed among English speaking children (Treiman, 1985, 1988), and, thus, casts doubt on the validity of the onset–rime dichotomy in explaining phoneme isolation performance in Arabic. Thus, it appears that, in Arabic, initial phoneme isolation is significantly more challenging for children than final phoneme isolation. This task is also more sensitive to variations in the linguistic affiliation of the target phoneme and of the stimulus word. Future research should address this issue directly using various types of phonemic awareness tasks that tap into the representation and processing of intrasyllabic units. It is possible that language-specific prosodic constraints may affect the ease with which various intrasyllabic units are orally produced as the outcome of phonological analysis. This may explain crosslinguistic variations in the ease with which initial versus final phonemes are isolated.

In summary, the results show that phoneme isolation is not an all-or-none phenomenon. Various aspects of the phonological and lexical representation of the word may affect the ease with which children can operate on phonemes. Previous work has shown that at least two major factors affect children's ability to perform mental operations on the phonological structure of words: the size of the phonological unit and the linguistic context (Treiman & Weatherson, 1992; Treiman & Zukowski, 1991). The results of the study show that at least two more factors are pertinent to explaining individual variations in phonological analysis in a diglossic context. These are the linguistic affiliation of the target phoneme and the lexical status of the stimulus word. Thus, in order to explain crosslinguistic variations in the route and the rate with which phonological awareness skills are acquired, current conceptualizations of the construct of phonemic awareness must be extended to incorporate other factors that are particularly pertinent to linguistic contexts characterized by linguistic disparity between the oral language of children and the written language.

APPENDIX A

CLUSTER I: INITIAL PHONEME ISOLATION

Spoken words

1. kami:l
2. ʕasal
3. ħali:b
4. ʈa:her
5. sa:ken
6. ma:lek
7. ʒama:l
8. daraʒ
9. ka:mel
10. ɣa:ʔem

Standard words

Spoken phonemes only	*Target phoneme standard*
1. ʃatam	1. θaman
2. s̲a:ʔed	2. θa:bet
3. χaʈar	3. θamar
4. da:ken	4. ðakar
5. ʈa:leb	5. ða:heb
6. sakab	6. ð̲alam
7. ʃa:meχ	7. ð̲a:fer
8. ʕadam	8. qa:reb
9. za:ʔel	9. qa:beʕ
10. ʒaʕal	10. qa:her

Pseudowords

Spoken phonemes only	*Target phoneme standard*
1. ʃakam	1. θamal
2. s̲a:beʕ	2. θa:meʈ
3. χabaʒ	3. θabal
4. da:ker	4. ðamal
5. ʈa:meʒ	5. ða:ħer
6. saħam	6. ð̲a:hel
7. ʃadel	7. ð̲aban
8. ʕafal	8. qa:lem
9. za:rem	9. qa:jer
10. ʒaħal	10. qafan

APPENDIX B

CLUSTER II: FINAL PHONEME ISOLATION

Spoken words

1. ʃa:reʕ
2. sa:ħer
3. ras̲i:f
4. ka:mel
5. ħabi:b
6. ʕazi:z
7. fa:hem
8. rabi:ʕ
9. sa:ken
10. χami:s

Standard words

Spoken phonemes only	*Target phoneme standard*
1. ba:reʕ	1. ʕa:beθ
2. zahi:d	2. wa:reθ
3. ʕami:l	3. wari:θ
4. s̲a:dem	4. mala:ð
5. ɣa:mer	5. nafað
6. ʒa:ħed	6. ba:heð̲
7. sana:m	7. ɣali:ð̲
8. suku:n	8. ʃa:heq
9. ʃa:meχ	9. ʈali:q
10. muhi:b	10. bari:q

Pseudowords

Spoken phonemes only	*Target phoneme standard*
1. sa:leʕ	1. na:beθ
2. maɣi:d	2. ħa:meθ
3. ʕafi:l	3. χani:θ
4. sa:rem	4. mari:ð
5. fa:ʕer	5. ma:leð
6. ħa:fed	6. ʕa:með̲
7. ʒana:m	7. ɣami:ð̲
8. ʃunu:n	8. sa:neq
9. sabi:n	9. dali:q
10. ʃa:deχ	10. sabi:q

Index to phonemic transcription

Symbol	*Description*
:	long vowel
θ	voiceless interdental fricative
ð	voiced interdental fricative
ð̲	emphatic voiced interdental fricative
ʃ	voiceless alveopalatal fricative
ʒ	voiced alveopalatal fricative
s̲	emphatic voiceless dental fricative
t̲	emphatic voiceless dental stop
d̲	emphatic voiced dental stop
χ	voiceless velar fricative
ɣ	voiced velar fricative
q	uvular stop
ʔ	glottal stop
ħ	voiceless pharyngeal fricative
ʕ	voiced pharyngeal fricative

ACKNOWLEDGMENTS

The author wishes to expresses her sincere gratitude to the anonymous reviewers for their invaluable feedback on an earlier version of the manuscript.

REFERENCES

Ababneh, Y. (2000). *Studies in the linguistics and phonology of Arabic*. Dar El-Shorook, Amman, Jordan (in Arabic).

Abu-Rabia, S. (2000). Effects of exposure to literary Arabic on reading comprehension in a diglossic situation. *Reading and Writing: An Interdisciplinary Journal, 13*, 147–157.

Adams, M. (1990). Beginning to read: Thinking and learning about print. Cambridge, MA: MIT Press.

Caravalos, M., & Bruck, M. (1993). The effect of oral and written language input in children's phonological awareness: A cross-linguistic study. *Journal of Experimental Child Psychology, 55*, 1–30.

Charity, A. H., Scarborough, H., & Griffin, D. M. (2002). *Familiarity with "School English" in African-American children and its relation to early reading achievement*. Paper presented at the annual meeting of the Society for the Scientific Study of Reading. Chicago.

Cossu, G., Shankweiler, D., Liberman, I. Y., Katz, L., & Tola, G. (1988). Awareness of phonological segments and reading ability in Italian children. *Applied Psycholinguistics, 9*, 1–16.

Durgunoglu, A. Y., & Oney, B. (1999). A cross-linguistic comparison of phonological awareness and word recognition. *Reading and Writing: An Interdisciplinary Journal, 11*, 281–299.

Elbro, C. (1996). Early linguistic abilities and reading development: A review and a hypothesis. *Reading and Writing: An Interdisciplinary Journal, 8*, 453–485.

Elbro, C. (1998). When reading is "readn" or somthn. Distinctness of phonological representations of lexical items in normal and disabled readers. *Scandinavian Journal of Psychology, 39*, 149–153.

Feitelson, D., Goldstein, Z., Iraqi, J., & Share, D. L. (1993). Effects of listening to story reading on aspects of literacy acquisition in a diglossic situation. *Reading Research Quarterly, 28*, 71–79.

Ferguson, C. A. (1959). Diglossia. *Word, 14*, 47–56.

Fledge, J. E. (1987). The production of "new" and "similar" phones in a foreign language: Evidence for the effect of equivalence classification. *Journal of Phonetics, 15*, 47–65.

Fowler, A. E. (1991). How early phonological development might set the stage for phoneme awareness. In S. Brady & D. Shankweiler (Eds.), *Phonological processes in literacy: A tribute to Isabelle Y. Liberman* (pp. 79–117). Hillsdale, NJ: Erlbaum.

Goswami, U., & Bryant, P. (1990). *Phonological skills and learning to read*. Hove, UK: Erlbaum.

Gough, P. B., Hoover, W. A., & Peterson, C. L. (1996). Some observations on a simple view of reading. In C. Cornoldi & J. Oakhill (Eds.), *Reading comprehension difficulties* (pp. 1–13). Mahwah, NJ: Erlbaum.

Gough, P. B., & Tunmer, W. E. (1986). Decoding, reading, and reading disability. *Remedial and Special Education, 7*, 6–10.

Hoover, W. A., & Gough, P. B. (1990). The simple view of reading. *Reading and Writing: An Interdisciplinary Journal, 2*, 127–160.

LaBerge, D., & Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology, 6*, 293–323.

Labov, W. (1995). Can reading failure be reversed: a linguistic approach to the question. In V. Gadsden & D. Wagner (Eds.), *Literacy among African-American youth: Issues in learning, teaching, and schooling* (pp. 39–68). Cresskill, NJ: Hampton Press.

Maamouri, M. (1998). *Language education and human development: Arabic diglossia and its impact on the quality of education in the Arab region,* Discussion paper prepared for The World Bank. The Mediterranean Development Forum.

Mann, V., & Wimmer, H. (2002). Phoneme awareness and pathways into literacy: A comparison of German and American children. *Reading & Writing: An Interdisciplinary Journal, 15*, 653–682.

McBride–Chang, C. (1995). What is phonological awareness? *Journal of Educational Psychology, 8*, 179–192.

Metsala, J. L. (1997a). An examination of word frequency and neighborhood density in the development of spoken word recognition. *Memory and Cognition, 25*, 47–56.

Metsala, J. L. (1997b). Spoken word recognition in reading disabled children. *Journal of Educational Psychology, 89*, 159–169.

Metsala, J. L. (1999). Young children's phonological awareness and non-word repetition as a function of vocabulary development. *Journal of Educational Psychology, 91*, 3–19.

Metsala, J. L., & Walley, A. C. (1998). Spoken vocabulary growth and the segmental restructuring of lexical representations: Precursors to phonemic awareness and early reading ability, In J. L. Metsala & L. C. Ehri (Eds.), *Word recognition in beginning literacy* (pp. 89–120). Mahwah, NJ: Erlbaum.

Oney, B., & Goldman, S. R. (1984). Decoding and comprehension skills in Turkish and English: Effects of the regularity of grapheme–phoneme correspondences. *Journal of Educational Psychology, 76*, 557–568.

Perfetti, C. A. (1985). *Reading ability*. New York: Oxford University Press.

Rickford, J. R., & Rickford, A. E. (1995). Dialect readers revisited. *Linguistics and Education, 7*, 107–128.

Saiegh–Haddad, E. (2003). Linguistic distance and initial reading acquisition: The case of Arabic diglossia. *Applied Psycholinguistics, 24*, 431–451.

Saiegh–Haddad, E. (2004). *Correlates of reading fluency in Arabic: Diglossic and orthographic factors*. Manuscript submitted for publication.

Treiman, R. (1985). Onset and rimes as units of spoken syllables: Evidence from children. *Journal of Experimental Child Psychology, 39*, 161–181.

Treiman, R. (1988). The internal structure of the syllable. In G. Carlson & M. Tanenhaus (Eds.), *Linguistic structure in language processing* (pp. 27–52). Norwell, MA: Kluwer Academic.

Treiman, R., & Weatherson, S. (1992). Effects of linguistic structure on children's ability to isolate initial consonants. *Journal of Educational Psychology, 84*, 174–201.

Treiman, R., & Zukowsky, A. (1991). Levels of phonological awareness. In S. A. Badry & D. P. Shankweiler (Eds.), *Phonological processes in literacy: A tribute to Isabelle Y. Liberman* (pp. 67–83). Hillsdale, NJ: Erlbaum.

Wade–Woolley, L., & Geva, E. (2000). Processing novel phonemic contrasts in the acquisition of L2 word reading. *Scientific Studies of Reading, 4*, 295–311.

Wagner, R. K., & Torgesen, J. K. (1987). The nature of phonological processing and its causal role in the acquisition of reading skills. *Psychological Bulletin, 101*, 192–212.

Walley, A. C., Metsala, J. L., & Garlock, V. M. (2003). Spoken vocabulary growth: Its role in the development of phoneme awareness and early reading ability. *Reading and Writing: An Interdisciplinary Journal, 16*, 5–20.

Walley, A. C., (1993). The role of vocabulary development in children's spoken word recognition and segmentation ability. *Developmental Review, 13*, 286–350.

Walley, A. C., & Fledge, J. E. (1999). Effects of lexical status on native and non-native vowel perception: A developmental study. *Journal of Phonetics, 27*, 307–332.

Wang, M., & Geva, E. (2003). Spelling acquisition of novel English phonemes in Chinese children. *Reading and Writing: An Interdisciplinary Journal, 16*, 325–348.

Washington, J. A., & Craig, H. K. (2001). Reading performance and dialectal variation. In J. L. Harris, A. G. Kamhi, & K. E. Pollock (Eds.), *Literacy in African American communities* (pp. 147–168). Mahwah, NJ: Erlbaum.

Wolf, M., & Katzir–Cohen, T. (2001). Reading fluency and its intervention. *Scientific Studies of Reading, 5*, 211–239.