Al Musawi *et al. Technol Innov Educ (2016) 2:12*
DOI 10.1186/s40660-016-0018-0

Technology, Innovation and Education

RESEARCH Open Access

The effectiveness of instructional software in reading comprehension skills and reading aloud of Omani fourth basic schools' students

Ali Al Musawi[1*], Ali Mahdi Kazem[2], Abdullah Al Hashmi[3] and Fatima Al Busaidi[3]

*Correspondence:
asmusawi@squ.edu.om
[1] Instructional and Learning Technologies Department, College of Education, Sultan Qaboos University, Muscat, Sultanate of Oman
Full list of author information is available at the end of the article

Abstract

This paper aims to describe the findings of a national research project which sought to evaluate the design and application of purpose built software for teaching reading in Arabic in the First Cycle of Basic Education (Grades 1–4) in Oman. Funded by his Majesty the Sultan of Oman's Grant for Strategic Research, an experimental methodology was used to investigate the effectiveness of e-learning on Grade 4 student outcomes in comprehension and reading aloud. The sample consisted of four fourth-class students' groups with 52 students for two experimental groups and 53 students for two control groups. A pre- and post-test was designed, reviewed and conducted. The research findings prove the effectiveness of the software on the students' achievement in reading comprehension; and that the students' attitudes towards the use of software in learning Arabic are positive. The research discusses these findings, suggests recommendations, and future steps.

Keywords: Software, Arabic reading, Language learning, Basic education, Oman

Background

A significant body of research supports the notion that best practice in the teaching and learning of language incorporates complementary use of technology with the development of foundational knowledge and use of language (Hoopingarner 2009; Goh et al. 2004). Sometimes referred to as e-language learning, myriad opportunities exist for the exploration of the innovative ways technology based strategies for language learning are being implemented (Spreen 2002).

With reference to Oman and the Arab World, Ismail et al. (2010) found that the use of technology by Arabic teachers for substantial learning and teaching activities was very modest and that use was predominantly confined to lesson preparation rather than a strategic teaching tool. Calls for the development of technological infrastructure, teacher and student skills and appropriate software which could be embedded into the curriculum by Al Musawi (2000) have been heeded as information and communication technology (ICT) and e-learning projects have become a key avenue for enhancing educational outcomes in Oman and the Arab World.

Research importance

Studies conducted in the field of reading have revealed that students in various stages suffer weakness in reading skills and understanding (Torgesen 2002). Al-Ramadhani (1995) reveals low performance of the third grade students at the Omani first cycle stage in oral reading skills and the prevalence of reading aloud errors specifically errors in the word structure, recognizing the end of words extension, and inversions. In addition, Al-Mawali (2003) also shows that post basic education students possess low possession of the reading aloud skills. The results of the Omani Ministry of Education (2005) shows that reading aloud errors are common among students at rates ranging between (91.7) and (65.7 %). Reading weakness among Omani students can be attributed to the spread of the vernacular in their daily lives, bilingualism, preservice teacher preparation, and poor school curriculum/textbooks design. More specifically, research emphasizes the lack of audio-visual aids and new technologies use in language teaching (To'eima 2004). However, the authors see that there is a lack of studies about integrating technology in Arabic language teaching specifically in the Omani context.

A project, financed by His Majesty, the Sultan of Oman's Grant for Strategic Research provided by Sultan Qaboos University Deanship of Research, was conducted to bridge this gap and show the extent to which educators, specifically Arabic language teachers, can be confident about the effectiveness of software technologies to improve student achievement in terms of reading skills and understanding. Results offer an important mechanism to justify, inspire, and strategically supplement interventions to address individual learning difficulties in the acquisition of Arabic language, particularly pertaining to reading comprehension.

Literature review

A review of the literature reveals that the use of e-language learning has led to higher performance in final examination results and increased the student abilities in the following areas: comprehension, analysis, dialogue, and initiative, along with its social, cultural and academic benefits. Other studies generally emphasize the positive returns of employing technologies in developing all language skills: listening, speaking, reading, and writing (Cunningham and Redmond 2008; Blake 1998; Beauvois 1997). However, the studies confirm that effective investment in technology extends beyond simply installing computers in the laboratory and is predicated upon well-designed software (Al Musawi 2000). Spreen (2002) addresses a number of design and implementation issues ranging from the recognition of learner dispositions, study habits, along with the appropriateness and range of activities. Culturally situated exercises and projects, authentic materials and experiences (especially audio and video), and methods that appeal to a variety of individual learning styles are also factors which influence engagement with e-language learning programs or experiences.

Katz and Carlisle (2009) research findings from the three case studies suggest that the participating fourth grade students did show growth in their reading and comprehension skills following instruction and practice with natural texts. Kim et al. (2010) study finds no significant difference between children in experimental and control groups' programs

on norm-referenced measures of word reading efficiency, reading comprehension, and vocabulary. They found a positive impact on oral reading fluency. Ertem (2010) study shows that there was a significant increase in students' reading comprehension when learning language through animated digital texts.

Before e-language learning rose to prominence, various studies identified key features of effective language teaching. In the field of Arabic language, an investigation into successful intervention strategies for elementary school reading and writing weaknesses, Al Najjar (1987) found that the social skills of the teacher were imperative instilling the confidence in the students to enable them to work independently. The results show that respect for individual difference, appreciation of specific learning difficulties and individual education plans assist in the development of language learning across all four domains: reading, writing, listening and speaking. Awadh (1999) studied the effectiveness of various therapeutic teaching strategies among female students-teachers relating to reading difficulties in preparatory schools. The results show statistical significant differences between the mean scores of female students-teachers in both the achievement and performance of the pre- and post-tests in favor of the post-test which confirms the effectiveness of the program in the cognitive as well as performance aspects. Al Barakat's (2010) conduct a study to investigate the effectiveness of a story strategy based training program on the development of Quranic assimilation of the basic schools' third-graders and their attitudes toward it. The results show that the experimental group achieved the highest means on the reading comprehension test; and show statistical significant differences between the two groups in favor of the experimental group. To detect the students' attitudes towards the training program, semi-structured interviewed were conducted. The results show positive attitudes and the story strategy appeared as the main tool in their preference towards reading. The most important factors that made them like it are: (1) design stories that are related to their environment; (2) employ the principle of the opportunity to retell the story; (3) discuss the narrative content; and (4) implement stimulating activities. The study recommends that the ministry of education employs the story strategy to develop reading comprehension of the basic schools' students; in addition to their teacher training on this strategy implementation. The study also calls for the need for further research by studying the relationship between developing reading comprehension skills among basic schools' students using the story using traditional method and computers through CD-ROMs to read the story in their homes and schools. Ertem suggests that comprehension by storytelling allows students with "to transform the story into their own words and also to share their individual understanding of text" (p. 148). Awadh (2010) finds that a proposed strategy in the treatment of reading, writing and achievement weakness in Arabic language for some basic education ninth grade students was effective in treating the weakness and recommends that call on the teachers to benefit from it and meet their students' differences through various curricular contents.

Al Musawi et al. (2014) notes the importance of developing computer software and word processing applications for local needs, which can then be adapted to the Arabic social environment and its curriculum in the field of Arabic language teaching. Dhir and Alsumait (2012) think that educational technology has revolutionized the traditional

forms of classroom teaching and learning. The paper examines and investigates the issue of providing appropriate educational technology and user interface services to Kuwaiti (and Arabic) students based on their needs and expectation. It aims to prepare a framework that acts as a guiding source for educationalists, teachers and policy makers for developing better educational technology and pedagogy services aiming young children. In a study conducted by Aldalalah et al. (2010), the impact of educational software on first elementary grade students in their learning of the Arabic language was investigated. In a comparison between students who used sounds with pictures and those who used text with pictures, language acquisition was greater in the former, thereby providing early indications of the merits of complementary technology with language learning. There are no significant differences from a statistical point of view due to gender or the interaction between gender and the application. Research into the impact of computerized educational games on fourth grade Jordanian students with difficulties in reading Arabic, revealed significant improvement in both direct and deferred achievement tests in favor of students who had been exposed computerized linguistic games (Al-Heela and Ghoneim 2002). Statistical significant differences were found between the mean scores students of study groups attributed to the gender in favor of females. Although the researchers found that improved outcomes were more significant in female than male students, other variables such as motivation and parental support needed to be taken into account. It was clear, however, that technology-enhanced learning has had a positive impact on language acquisition (Al-Khatib 2011).

In Oman, studies by Al-Ramadhani (1995) and Al-Fori (1999) revealed weaknesses in both silent and oral reading skills as well as comprehension in Arabic among students in different grades. A study conducted by the Ministry of Education (2005) on the common errors of reading aloud among the Omani first cycle basic education schools' students (such as: letter additions, deletions, inversions, and repetitions) found a high rate of errors at a ratio of (91.7 %). In an analysis of oral reading skills, Al-Mawali (2003) noted an error rate of (41.8 %) with the visual recognition of printed symbols (letters, words, and sentences) as the most mastered skills were at (61.2 %). Based on these results, Al-Mawali emphasized the importance of developing teaching methods which exploit the range of individualized, targeted learning experiences using an online platform. Other studies (Al-Hashmi 2009; Al-Yazidi 2006) also advocate for the efficacy of e-language learning, particularly for the development of reading and the advantages of allocating multimedia classrooms with teachers skilled to exploit the use of technology as a strategy inherent in the Arabic language acquisition.

It can be said that past research (To'eima 2004; Almekhlafi and Almeqdadi 2010; Ismail et al. 2010) used surveys to explore the teachers' perceptions on using instructional software. Further, there is a lack of studies about integrating such software in Arabic language teaching specifically in the Omani context (see for example Al-Ramadhani 1995; Al-Fori 1999).

Research objectives

Notwithstanding the effectiveness of using instructional software in teaching reading, it seems that the majority of the reviewed studies lack experimental studies focused on

measuring the effectiveness of this software on their achievement and comprehension of reading Arabic language. This paper attempts to provide empirical evidence on the use of instructional software and its impact and improvement of Omani students' Arabic reading comprehension and skills in classroom environment. Moreover, it covers the lack of design, development, implementation and evaluation of software tailored to the needs of these students in Arabic language learning (see for example Aldalalah et al. 2010; Almekhlafi and Almeqdadi 2010; Al Musawi 2000).

In light of the above shortage in the literature, this research aims to describe the findings of a national research project that investigates the effectiveness of the design, production, use and evaluation of new software for reading in Arabic in Grades 1–4 which comprise the First Cycle of Basic Education in Oman. The following reading weaknesses are studied in this research:

1. Reading comprehension weaknesses items: understand content, realize the temporal and locative arrangement, distinguish between compact and pronounced article 'lam', recognize end of words extension 'med' and formation 'tanween', differentiate between folded and normal letter 'taa', and reading comprehension.
2. Reading aloud weaknesses items: additions, deletions, inversions, repetitions, intermittent reading, grammar error, morphological error, erroneous stop, and reading aloud.

Problem, questions and hypotheses

The overarching research question is "what is the effect of this new software on student achievement in reading comprehension". More specifically, the research poses the following questions based on the results of students using the software and the control group which does not:

1. Does student achievement differ in terms of reading comprehension skills (understand content, realize the temporal and locative arrangement, distinguish between compact and pronounced article 'lam', recognize end of words extension 'med' and formation 'tanween', differentiate between folded and normal letter 'taa', reading comprehension)?

   - **Hypothesis 1**: There are no statistically significant differences at the level of <0.05 between experimental and control groups in the post-test in reading comprehension.
   - **Hypothesis 2**: There are no statistically significant differences at the level of <0.05 between pre- and post-tests in reading comprehension for the experimental group.

2. Does student achievement differ in terms of oral reading skills (additions, deletions, inversions, repetitions, intermittent reading, grammar error, morphological error, erroneous stop, reading aloud)?

   - **Hypothesis 3**: There are no statistically significant differences at the level of <0.05 between experimental and control groups in the post-test in reading aloud errors.

- **Hypothesis 4**: There are no statistically significant differences at the level of <0.05 between pre- and post-tests in reading aloud errors for the experimental group.

3. Do attitudes about the use of software in Arabic language acquisition change with experience of such software?

- **Hypothesis 5**: There are no statistical significant differences at the level of <0.05 between experimental and control groups in the post-test in the students' attitudes towards the use of software in learning Arabic.
- **Hypothesis 6**: *There are no statistical significant differences at the level of <0.05* between pre- and post-tests in the students' attitudes towards the use of software in learning Arabic for the experimental group.

Methods

The research project follows a quasi-experimental approach. As stated above, the past research lacks (1) the empirical evidence on the use of instructional software and its impact on Omani students' reading comprehension and skills in classroom environment (To'eima 2004; Almekhlafi and Almeqdadi 2010; Ismail et al. 2010); and (2) the design, development, implementation and evaluation of such a software in Arabic language learning (Almekhlafi and Almeqdadi 2010; Al-Hashmi 2009; Al-Yazidi 2006; Al Musawi 2000).

Research design

The research follows the following four phases:

Phase I—Analysis and design: in this phase, the literature linking education technology and its applications in teaching Arabic language is surveyed, with a focus on the reading weaknesses. The design, production and use standards of Arabic language software that address the reading weaknesses are derived from the literature. To address these weaknesses, Al-Busaidi et al. (2013) proposes standards for the design of linguistic software derived from the reviewed research. In addition, the works of Al Barakat (2010), Ertem (2010) and Awadh (2010) that recommended the use of storytelling strategy were used in the design processes.

Fourth class reading curricular content with specific reading skills for software production was then selected; and linguists and teachers in cooperation with instructional and multimedia designers prepare scripts and designs requirements to produce the software.

Phase II—Development: in this phase, content of the software was measured and prepared in its final form. Further, the software was produced using different animation and multimedia tools (see software screen shot below). It was reviewed in line with the design standards by experts and specialists.

Phase III—Implementation: in this phase, the experiment was conducted, new software was implemented, and the pre- and post-tests applied in a timely fashion. The following design was implemented for each of the three fourth-class students' groups:

Group	Pre-tests	Experimental treatment	Post-tests
Experimental group	Achievement test (paper)	Software use	Achievement test (paper)
Control group	Achievement test (paper)	Use traditional method of teaching (without software)	Achievement test (paper)
Experimental group	O_1	X	O_2
Control group	O_1		O_2

The implementation of the experiment took place in two educational regions schools (Muscat and South Batinah) in collaboration with teachers and specialists using their learning resource centers. The implementation process took one-month time. Data were then collected and analyzed.

Population and sample

The study population consists of all the students of the first cycle of the basic education schools in two Omani educational regions with a randomly selected samples consisted of four fourth-class students' groups with 52 students for two experimental groups and 53 students for two control groups. The students' are males and females' within 9 years old category and different levels of achievement in reading. The study glean from the experimental design of Kim et al. (2010), Awadh (1999), Al Barakat's (2010), Aldalalah et al. (2010), and Katz and Carlisle (2009) to reinforce its sampling choice.

Instruments, validity and reliability

Based on previous studies of Kim et al. (2010), Awadh (1999), Al Barakat's (2010), Aldalalah et al. (2010), and Katz and Carlisle (2009), an achievement test to measure reading comprehension and oral reading was devised to be administered both prior to, and after the implementation of the software as a learning tool. The face validity of the test was ascertained after scrutiny by a panel of referees, all of whom were specialists in Arabic language acquisition and assessment. On the basis of the panel's suggestions, the test was refined to produce the final version. The reliability coefficient was measured using Cronbach's alpha ($\alpha = 0.73$). In addition, an attitudinal scale was also developed to measure the student attitudes towards using the software and it too was reviewed by a group of referees and modified accordingly to produce a reliability coefficient using Cronbach's alpha ($\alpha = 0.77$) (refer to Appendix: Table 17).

Methods of analysis

Considering the type of variables, the study uses the statistical program (SPSS) to calculate means, standard deviations, t test and analysis of variance (MANOVA, ANCOVA) and to immediately provide the researchers with relevant tests/results.

Groups equivalence

Reading comprehension equivalence

To ensure the groups' equivalence, means and standard deviations are calculated for both the experimental and control groups' pre-test scores in reading comprehension. Table 1 includes them.

Table 1 Means (M) and standard deviations (SD) of the experimental and control groups pre-test scores in reading comprehension

Read. Comp. Items	Group	*N*	*M*	SD
Understand content	Experimental	52	0.38	0.28
	Control	53	0.35	0.26
Realize the temporal and locative arrangement	Experimental	52	0.26	0.31
	Control	53	0.22	0.25
Distinguish between compact and pronounced article 'lam'	Experimental	52	0.78	0.38
	Control	53	0.71	0.31
Recognize end of words extension 'med' and formation 'tanween'	Experimental	52	0.41	0.30
	Control	53	0.24	0.24
Differentiate between folded and normal letter 'Taa'	Experimental	52	0.52	0.37
	Control	53	0.56	0.31
Reading comprehension	Experimental	52	0.49	0.17
	Control	53	0.41	0.16

To examine the difference significance of the means, multivariate analysis of variance (MANOVA) is used. Analysis shows that Wilks' Lambda $= 0.89$, $F = 2.55$, $p = 0.032$, at $dF = 101{,}5$ which indicates that differences exist between the experimental and control groups in some or all items. Table 2 summarizes MANOVA results.

Table 2 shows that F value is significant in two items (recognize end of words extension 'med' and formation 'tanween'; and reading comprehension in general) where the means of the experimental group are higher than those of the control group.

Reading aloud equivalence

To ensure the groups' equivalence, means and standard deviations are calculated for both the experimental and control groups pre-test scores in reading aloud errors. Table 3 includes them.

To examine the difference significance of the means, multivariate analysis of variance (MANOVA) is used. Analysis shows that Wilks' Lambda $= 0.70$, $F = 5.36$, $p < 0.001$, at $dF = 96{,}8$ indicating that differences exist between the experimental and control groups in some or all items. Table 4 summarizes MANOVA results.

Table 2 Summary of multivariate analysis of variance (MANOVA) comparing the experimental and control groups pre-test in reading comprehension

Source	Read. Comp. Items	SS	MS	$F_{(1, 103)}$	*p*
Group	Understand content	0.029	0.029	0.409	0.524
	Realize the temporal and locative arrangement	0.057	0.057	0.722	0.397
	Distinguish between compact and pronounced article 'lam'	0.142	0.142	1.181	0.280
	Recognize end of words extension 'med' and formation 'tanween'	0.751	0.751	10.250	0.002
	Differentiate between folded and normal letter 'Taa'	0.042	0.042	0.361	0.549
	Reading comprehension	0.155	0.155	5.747	0.018
Error	Understand content	7.512	0.072		
	Realize the temporal and locative arrangement	8.309	0.079		
	Distinguish between compact and pronounced article 'lam'	12.594	0.120		
	Recognize end of words extension 'med' and formation 'tanween'	7.692	0.073		
	Differentiate between folded and normal letter 'Taa'	12.185	0.116		
	Reading comprehension	2.831	0.027		

Table 3 Means (M) and standard deviations (SD) of the experimental and control groups pre-test scores in reading aloud

Read. Aloud Items	Group	*N*	*M*	SD
Additions	Experimental	52	5.54	3.41
	Control	53	4.56	4.64
Deletions	Experimental	52	5.10	3.33
	Control	53	6.73	6.39
Inversions	Experimental	52	5.44	4.00
	Control	53	6.07	7.34
Repetitions	Experimental	52	4.60	3.39
	Control	53	5.31	3.85
Intermittent reading	Experimental	52	6.46	4.66
	Control	53	10.42	9.14
Grammar error	Experimental	52	14.17	7.01
	Control	53	12.31	7.83
Morphological error	Experimental	52	3.48	3.03
	Control	53	6.42	5.95
Erroneous stop	Experimental	52	4.65	4.21
	Control	53	5.00	4.26
Reading aloud	Experimental	52	49.44	20.03
	Control	53	56.82	33.26

Table 4 Summary of multivariate analysis of variance (MANOVA) comparing the experimental and control groups pre-test in reading aloud errors

Source	Read. Aloud Items	SS	MS	$F_{(1, 103)}$	*p*
Group	Additions	25.40	25.40	1.52	0.220
	Deletions	71.11	71.11	2.69	0.104
	Inversions	10.62	10.62	0.30	0.585
	Repetitions	13.59	13.59	1.03	0.313
	Intermittent reading	418.44	418.44	7.82	0.006
	Grammar error	92.87	92.87	1.68	0.198
	Morphological error	230.63	230.63	10.17	0.002
	Erroneous stop	3.20	3.20	0.18	0.674
	Reading aloud	1454.15	1454.15	1.90	0.171
Error	Additions	1754.45	16.71		
	Deletions	2773.43	26.41		
	Inversions	3726.54	35.49		
	Repetitions	1388.27	13.22		
	Intermittent reading	5618.31	53.51		
	Grammar error	5815.19	55.38		
	Morphological error	2380.36	22.67		
	Erroneous stop	1885.77	17.96		
	Reading aloud	80189.01	763.71		

Table 4 shows that *F* value is significant in two items (intermittent reading; and morphological error) where the means of the control group are higher than those of the experimental group.

Equivalence in attitudes towards the use of software in learning Arabic

To ensure the groups' equivalence, a *t* test comparing the experimental and control groups pre-test in attitudes towards the use of software in learning Arabic, is highlighted in Table 5.

Table 5 shows that *t* value is statistically significant where the means of the experimental group are higher than those of the control group.

In the following sections, findings from the research data are presented, analyzed, and discussed.

Results and findings

Hypothesis 1 "There are no statistical significant differences at the level of <0.05 between experimental and control groups in the post-test in reading comprehension". To test this hypothesis, arithmetic means and standard deviations of the experimental and control groups scores are calculated. Table 6 includes them.

To examine the significance of the difference in the means, multivariate analysis of variance (MANOVA) was used. Analysis shows that Wilks' Lambda = 0.36, $F = 35.39$, $p < 0.001$, at $dF = 101{,}5$ with effect size of (0.64) suggesting a very large effect size according to Cohen's criteria. This finding indicates that employing this software produced an effect on the experimental groups' achievement in terms of reading comprehension. Table 7 summarizes MANOVA results.

Table 5 Results of *t* test of the experimental and control groups pre-test in attitudes towards the use of software in learning Arabic

Group	*N*	*M*	SD	*t*	*p*
Experimental	52	35.29	0.51	3.76	0.001
Control	52	32.71	0.45		

Table 6 Means (M) and standard deviations (SD) of the experimental and control groups post-test scores in reading comprehension

Read. Comp. Items	Group	*N*	*M*	SD
Understand content	Experimental	52	0.70	0.27
	Control	53	0.34	0.29
Realize the temporal and locative arrangement	Experimental	52	0.61	0.38
	Control	53	0.21	0.24
Distinguish between compact and pronounced article 'lam'	Experimental	52	0.97	0.11
	Control	53	0.65	0.36
Recognize end of words extension 'med' and formation 'tanween'	Experimental	52	0.80	0.19
	Control	53	0.33	0.29
Differentiate between folded and normal letter 'Taa'	Experimental	52	0.78	0.30
	Control	53	0.61	0.34
Reading comprehension	Experimental	52	0.80	0.14
	Control	53	0.43	0.18

Table 7 Summary of multivariate analysis of variance (MANOVA) comparing the experimental and control groups post-test in reading comprehension

Source	Read. Comp. Items	SS	MS	$F_{(1, 103)}$	p	Effect size
Group	Understand content	3.45	3.45	43.20	<0.001	0.29
	Realize the temporal and locative arrangement	4.11	4.11	41.22	<0.001	0.28
	Distinguish between compact and pronounced article 'lam'	2.89	2.89	40.05	<0.001	0.28
	Recognize end of words extension 'med' and formation 'tanween'	5.93	5.93	96.48	<0.001	0.48
	Differentiate between folded and normal letter 'Taa'	0.73	0.73	6.93	<0.001	0.06
	Reading comprehension	3.50	3.50	127.44	<0.001	0.55
Error	Understand content	8.39	0.08			
	Realize the temporal and locative arrangement	10.47	0.10			
	Distinguish between compact and pronounced article 'lam'	7.58	0.07			
	Recognize end of words extension 'med' and formation 'tanween'	6.46	0.06			
	Differentiate between folded and normal letter 'Taa'	11.06	0.11			
	Reading comprehension	2.89	0.03			

Both Table 7 and Fig. 1 show that the F value is significant for all reading comprehension skills. Compared with data shown in Table 1, this finding shows that the intervention was an effective way to improve test scores.

Hypothesis 2 "There are no statistical significant differences at the level of <0.05 between experimental and control groups in the post-test in reading aloud errors". To test this hypothesis, arithmetic means and standard deviations of the experimental and control groups scores are calculated. Table 8 includes them.

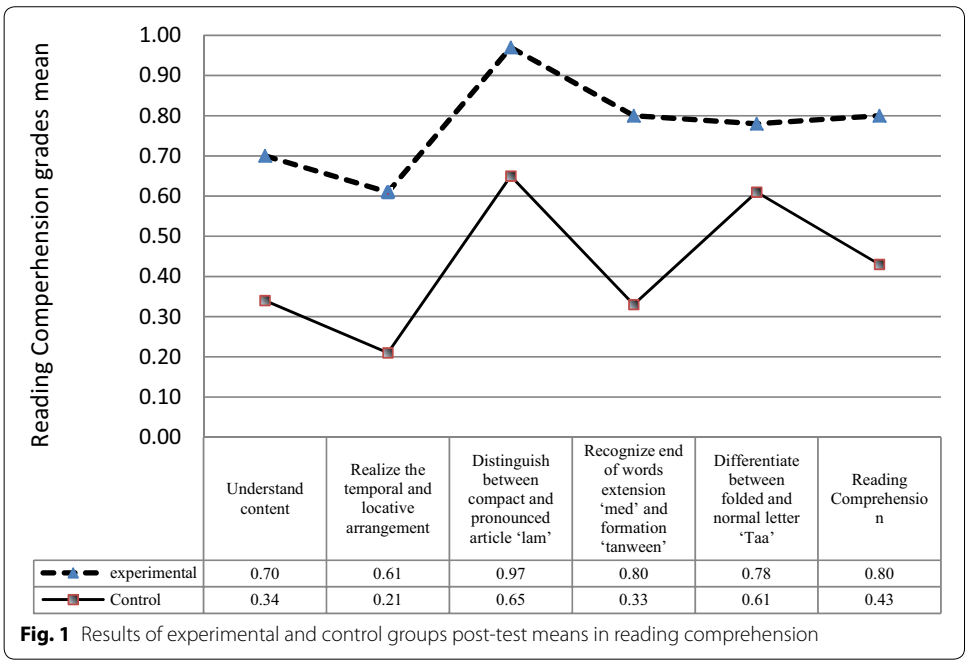

Fig. 1 Results of experimental and control groups post-test means in reading comprehension

Table 8 Means and standard deviations of the experimental and control groups post-test scores in reading aloud

Read. Aloud items	Group	*N*	*M*	SD
Additions	Experimental	52	3.04	2.01
	Control	53	2.51	2.15
Deletions	Experimental	52	1.69	1.57
	Control	53	5.21	8.97
Inversions	Experimental	52	2.12	1.83
	Control	53	2.34	2.10
Repetitions	Experimental	52	2.38	2.34
	Control	53	2.66	2.11
Intermittent reading	Experimental	52	2.98	3.68
	Control	53	4.36	4.04
Grammar error	Experimental	52	8.92	5.04
	Control	53	7.83	6.39
Morphological error	Experimental	52	1.12	1.13
	Control	53	2.40	2.11
Erroneous stop	Experimental	52	1.17	1.49
	Control	53	2.58	2.86
Reading aloud	Experimental	52	23.42	11.83
	Control	53	29.89	18.94

To examine the significance of the different means, multivariate analysis of variance (MANOVA) is used. Analysis shows that Wilks' Lambda = 0.72, $F = 4.57$, $p < 0.001$, at $dF = 96,8$ with effect size of (0.28) suggesting a very large effect size according to Cohen criteria. This finding indicates that employing this software is effective in terms of minimizing reading aloud errors of the experimental groups. Table 9 summarizes MANOVA results. This finding affirms that the group using the software demonstrated significantly less reading aloud errors than the control group.

Both Table 9 and Fig. 2 show that F value of deletions, morphological errors, erroneous stop, and the overall grade of reading aloud is significant, whereas other items are not.

Hypothesis 3 "There are no statistically significant differences at the level of <0.05 between experimental and control groups in the post-test in the students' attitudes towards the use of software in learning Arabic".

To test this hypothesis, arithmetic means and standard deviations of the experimental and control groups scores in terms of the attitudes towards the use of software in learning Arabic are calculated. Table 10 includes them.

To examine the difference significance of the means, analysis of covariance (ANCOVA) is used. Table 11 includes them.

Analysis in Table 11 shows that the group and results on the pre-test have statistical significant effect at the level of <0.001 and 0.005 consecutively; suggesting a very large effect size according to Cohen's criteria. To know the group's differences direction, means are referred to (due to the statistical significance of the pre-test); where the experimental group's mean is 36.12 and the control group's mean is 31.98. This difference

Table 9 Summary of multivariate analysis of variance (MANOVA) comparing the experimental and control groups post-test in reading aloud

Source	Read. Aloud items	SS	MS	$F_{(1, 103)}$	*p*	Effect size
Group	Additions	7.35	7.35	1.70	0.195	–
	Deletions	324.34	324.34	7.76	0.006	0.07
	Inversions	1.32	1.32	0.34	0.562	–
	Repetitions	2.00	2.00	0.40	0.528	–
	Intermittent reading	49.82	49.82	3.34	0.071	–
	Grammar error	31.35	31.35	0.94	0.333	–
	Morphological error	43.06	43.06	14.99	<0.001	0.13
	Erroneous stop	52.32	52.32	10.01	0.002	0.09
	Reading aloud	1096.62	1096.62	4.38	0.039	0.04
Error	Additions	445.17	4.32			
	Deletions	4307.79	41.82			
	Inversions	401.19	3.90			
	Repetitions	512.19	4.97			
	Intermittent reading	1537.17	14.92			
	Grammar error	3419.16	33.20			
	Morphological error	295.99	2.87			
	Erroneous stop	538.31	5.23			
	Reading aloud	25780.01	250.29			

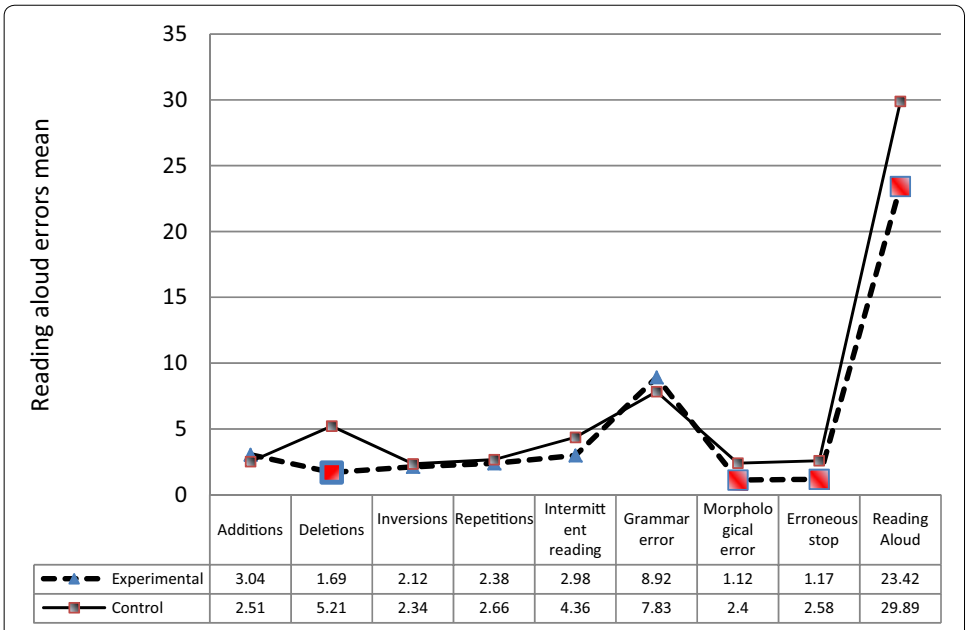

Fig. 2 Results of experimental and control groups post-test in reading aloud

Table 10 Means and standard deviations of the experimental and control groups post-test scores in attitudes towards using software in learning Arabic

Group	*N*	*M*	SD
Experimental	52	36.44	1.99
Control	51	31.65	4.21

Table 11 Summary of analysis of covariance (ANCOVA) of group effects in post-test scores in attitudes towards using software in learning Arabic

Source	SS	MS	$F_{(100,1)}$	*p*	Effect size
Pre-test	84.74	84.74	8.43	0.005	0.08
Group	388.90	388.90	38.67	<0.001	0.28
Error	1005.74	10.06			

between the two groups indicates that the experimental group developed stronger, more positive attitudes about the use of software than the control group. This finding suggests that employing this software is produced a positive effect in terms of reading comprehension and the concomitant results on the tests. Accordingly, the null hypothesis is refuted and the alternate hypothesis, which indicate that statistical significant differences between both groups in post-test of the attitudes towards using software in learning Arabic exist at the level of <0.005 in favor of the experimental group, is accepted.

Hypothesis 4 "There are no statistically significant differences at the level of <0.05 between pre- and post-tests in reading comprehension for the experimental group". To test this hypothesis, arithmetic means and standard deviations of the experimental and control groups scores are calculated. Table 12 includes them.

To examine the significance of the difference in the means, multivariate analysis of variance (MANOVA) is used. Analysis shows that Wilks' Lambda = 0.15, $F = 53.62$, $p < 0.001$, at $dF = 47{,}5$ with effect size of (0.85) suggesting a very large effect size according to Cohen criteria. This finding indicates that employing this software is effective in terms of post-test results as clear improvement is observed among the experimental group's students in terms of reading comprehension. Table 13 summarizes MANOVA results of experimental group treatment in reading comprehension.

Table 12 Means and standard deviations of the experimental and control groups pre- and post-test scores in reading comprehension ($n = 52$)

Read. Comp. Items	Measure	Mean	SD
Understand content	Pre-test	0.38	0.28
	Post-test	0.70	0.27
Realize the temporal and locative arrangement	Pre-test	0.26	0.31
	Post-test	0.61	0.38
Distinguish between compact and pronounced article 'lam'	Pre-test	0.78	0.38
	Post-test	0.97	0.11
Recognize end of words extension 'med' and formation 'tanween'	Pre-test	0.41	0.30
	Post-test	0.80	0.19
Differentiate between folded and normal letter 'Taa'	Pre-test	0.52	0.37
	Post-test	0.78	0.30
Reading comprehension	Pre-test	0.49	0.17
	Post-test	0.80	0.14

Table 13 Summary of multivariate analysis of variance (MANOVA) of experimental group treatment in reading comprehension

Source	Read. Comp. Items	SS	MS	$F_{(51,1)}$	p	Effect size
Group	Understand content	2.57	2.57	32.34	<0.001	0.39
	Realize the temporal and locative arrangement	3.03	3.03	22.67	<0.001	0.31
	Distinguish between compact and pronounced article 'lam'	0.99	0.99	15.37	<0.001	0.23
	Recognize end of words extension 'med' and formation 'tanween'	4.09	4.09	99.10	<0.001	0.66
	Differentiate between folded and normal letter 'Taa'	1.69	1.69	23.55	<0.001	0.32
	Reading comprehension	2.45	2.45	166.92	<0.001	0.77
Error	Understand content	4.05	0.08			
	Realize the temporal and locative arrangement	6.81	0.13			
	Distinguish between compact and pronounced article 'lam'	3.30	0.07			
	Recognize end of words extension 'med' and formation 'tanween'	2.11	0.04			
	Differentiate between folded and normal letter 'Taa'	3.66	0.07			
	Reading comprehension	0.75	0.02			

Table 13 shows that the software has an effect on all reading comprehension items which indicates its effectiveness. Figure 3 also shows the improvement occurred to the reading comprehension skills of the experimental group's students.

Hypothesis 5 "There are no statistical significant differences at the level of <0.05 between pre- and post-tests in reading aloud errors for the experimental group". To test this hypothesis, arithmetic means and standard deviations of the experimental and control groups scores are calculated. Table 14 includes them.

To examine the difference significance of the means between the pre- and post-test of the reading aloud, multivariate analysis of variance (MANOVA) is used. Analysis shows that

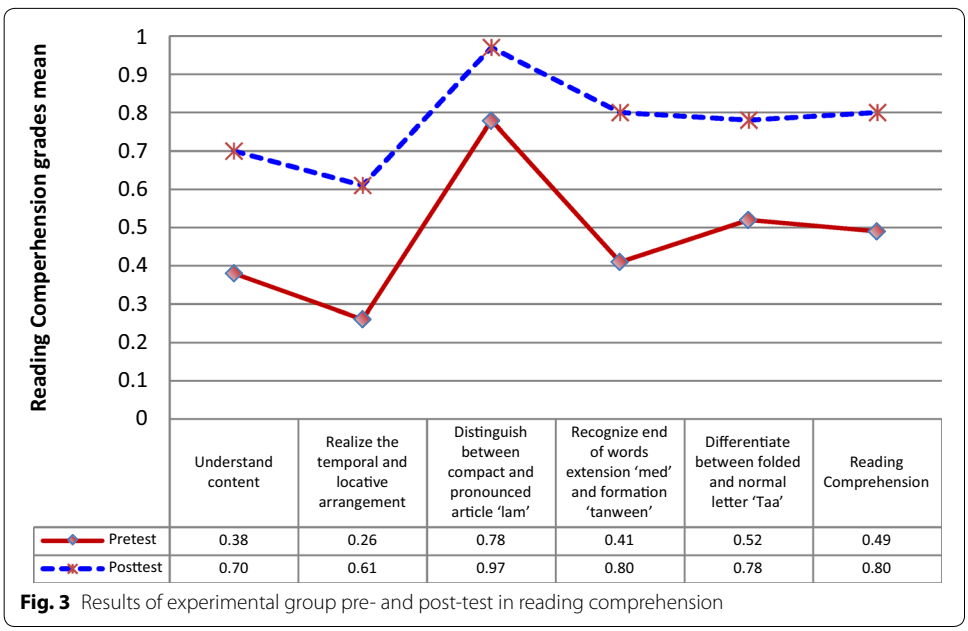

Fig. 3 Results of experimental group pre- and post-test in reading comprehension

Table 14 Means and standard deviations of the experimental and control groups pre- and post-test scores in reading aloud ($n = 52$)

Read. Comp. Items	Measure	Mean	SD
Additions	Pre-test	5.54	3.41
	Post-test	3.04	2.01
Deletions	Pre-test	5.10	3.33
	Post-test	1.69	1.57
Inversions	Pre-test	5.44	4.00
	Post-test	2.12	1.83
Repetitions	Pre-test	4.60	3.39
	Post-test	2.38	2.34
Intermittent reading	Pre-test	6.46	4.66
	Post-test	2.98	3.68
Grammar error	Pre-test	14.17	7.01
	Post-test	8.92	5.04
Morphological error	Pre-test	3.48	3.03
	Post-test	1.12	1.13
Erroneous stop	Pre-test	4.65	4.21
	Post-test	1.17	1.49
Reading aloud	Pre-test	49.44	20.03
	Post-test	23.42	11.83

Wilks' Lambda = 0.108, $F = 45.21$, $p < 0.001$, at $dF = 44,8$ with effect size of (0.89) suggesting a very large effect size according to Cohen criteria. This finding indicates that employing this software has post-test results as clear reduction of reading aloud errors is observed among the experimental group's students. Table 15 summarizes MANOVA results.

Table 15 shows that the software has an effect on all reading aloud errors items which indicates its effectiveness. Figure 4 also shows the improvement occurred in the reduction of reading aloud errors of the experimental group's students.

Hypothesis 6 "There are no statistically significant differences at the level of <0.05 between pre- and post-tests in the students' attitudes towards the use of software in learning Arabic for the experimental group". To test this hypothesis, arithmetic means and standard deviations of the experimental group's attitudes in pre- and post-test towards using software in learning Arabic are calculated along with using *t* tests of correlated samples. See Table 16.

Analysis in Table 16 shows that the *t* test value is 0.046 indicating a statistically significant effect according to Cohen's criteria. To know the group's differences direction, means are referred to; where the post-test mean is 36.44 which is larger than the pre-test mean of 35.29 with a difference of 1.12. This difference between both tests indicates that employing this software has a positive effect on improving the experimental group's students' attitudes towards using software in learning Arabic. Accordingly, the null hypothesis is refuted and the alternate hypothesis, which supports a statistically significant difference between the pre- and post-test scores of the experimental group attitudes towards using software in learning Arabic exist at the level of <0.005 in favor of the post-test, is accepted.

Table 15 Summary of multivariate analysis of variance (MANOVA) of experimental group treatment in reading aloud

Source	Read. Aloud items	SS	MS	$F_{(51,1)}$	*p*	Effect size
Group	Additions	1912.65	1912.65	234.29	<0.001	0.821
	Deletions	1198.16	1198.16	129.10	<0.001	0.717
	Inversions	1485.09	1485.09	110.83	<0.001	0.685
	Repetitions	1267.01	1267.01	98.28	<0.001	0.658
	Intermittent reading	2318.09	2318.09	106.56	<0.001	0.676
	Grammar error	13,869.24	13,869.24	271.71	<0.001	0.842
	Morphological error	549.24	549.24	88.29	<0.001	0.634
	Erroneous stop	882.78	882.78	73.60	<0.001	0.591
	Reading aloud	138,043.47	138,043.47	341.63	<0.001	0.870
Error	Additions	416.35	8.16			
	Deletions	473.34	9.28			
	Inversions	683.41	13.40			
	Repetitions	657.49	12.89			
	Intermittent reading	1109.41	21.75			
	Grammar error	2603.26	51.04			
	Morphological error	317.26	6.22			
	Erroneous stop	611.72	12.00			
	Reading aloud	20,608.03	404.08			

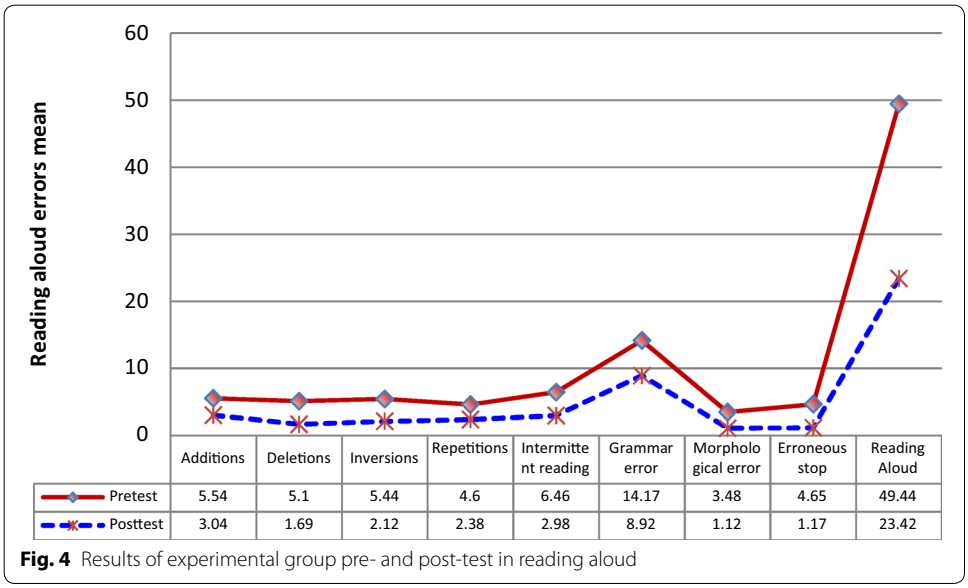

Fig. 4 Results of experimental group pre- and post-test in reading aloud

Table 16 Results of *t* test of correlated samples for the effect of pre- and post-test scores towards using software in learning Arabic (*n* = 52)

Test	Mean	SD	*t*	*p*	Effect size
Pre-test	35.29	3.71	2.04	0.046	0.08
Post-test	36.44	1.99			

Discussion of the findings

This paper aims to describe the design, implementation, and findings of a national scale research project that investigated the effectiveness of the design and use of new software in learning Arabic reading at Omani first basic education cycle's schools.

Findings indicate that the experimental group students produced significantly better post-test scores and this success was mirrored by their positive attitudes to the use of the software. *F* value is significant for all reading comprehension skills. Improvements were registered by the experimental group on all reading comprehension items which suggests the software effectiveness. Findings also show the improvement occurred to the reading comprehension skills of the experimental group's students. These findings are corroborated by Spreen (2002), Ertem (2010), Al-Khatib (2011), Cunningham and Redmond (2008). Findings indicate that the experimental group exceeded the control group in terms of their preference for using software in the learning of Arabic. Thus, the software is a major contributing factor in the improvement of reading comprehension skills in Arabic, as well minimizing reading aloud errors. F values of deletions, morphological errors, erroneous stop, and the overall grade of reading aloud were significant. These findings also corroborate the results of previous studies by Blake (1998), Beauvois (1997), Al Musawi (2000), Al-Mawali (2003), Al-Hashmi (2009) and Al-Yazidi (2006).

The scores of students using this software on post-tests revealed significantly better results than those students in the control group. Across the range of comprehension and oral reading variables, students in the experimental group out-performed their peers in the control group. It may also be inferred that increased success explained the positive correlation with experimental group attitudes to the use of software in Arabic language acquisition, thereby corroborating findings by Ismail et al. (2010) and Aldalalah et al. (2010). On the contrary, the above results may refute findings of Judson (2010) which shows no significant improvements in students' reading skills due to technology use. They also disagree with OECD (2015) which reports no noticeable improvement is attributed to technology use in terms of students' results for reading (p. 3).

Limitations and future research/steps

Limitations of this research project could be summarized seen in small size of the sample from two basic education schools; subject expectancy which normally affects this type of experimental research; and that reliability of the instrument used in this study will perhaps need further validation. However, those limitations did not considerably influence the study findings.

The following steps are considered for future implementation:

1. Conduct a training workshop for the Arabic language teachers to investigate their perceptions of using the software in their teaching. Interviews will be conducted and the results of this research will be reported in an article.
2. The authors start procedures for the patentability of the new software through the Academic Innovation Assistance Program at Sultan Qaboos University. They register their interests in applications for patent and record all related documents.

Future research may include further investigation of the software effects on other Arabic language skills such as speaking and writing and its effect on different categories of students such as those with higher levels of reading difficulties.

Study implications

The study results are in tandem with the Omani Ministry of Education efforts for effective implementation of multimedia technologies and software in the instruction and learning of Arabic language. The software used in this study specifically addresses the reading comprehension and weaknesses among Omani students. Through practical implementation and experimental treatment with fourth class graders, it was proved that the use of this software improved the students' achievement in acquiring reading skills. The software allows the students to visualize the reading topics using storytelling instructional approach and practice their reading to support their retention and comprehension of information.

Conclusions and recommendations

This research affirms the effectiveness of the software design and its positive influence on student achievement in comprehending and reading Arabic aloud. Student success correlated with positive attitudes to the use of software when learning Arabic. Results of the study suggest the potential to significantly reduce reading errors and enhance comprehension by integrating technology into Arabic language classrooms in Oman. In sum, data and findings in this paper support the idea that the intervention was an effective way to improve test scores and that it is an improvement over nonintervention-based instruction.

The research recommends that technology based applications may be used to improve language skills achievement and eradicate weaknesses among students in the early stages of their schooling. Students' attitudes show that using technology has generally a positive effect on improving their language learning, specifically Arabic in this context. However, the design and implementation of technology should be informed by Arabic language excellence and culturally appropriate standards. Teachers, educationalists, and stakeholders can use these findings and software production standards to design, develop, use, and evaluate technological application in Arabic language teaching.

Author details
[1] Instructional and Learning Technologies Department, College of Education, Sultan Qaboos University, Muscat, Sultanate of Oman. [2] Psychology Department, College of Education, Sultan Qaboos University, Muscat, Sultanate of Oman. [3] Curriculum and Instruction Department, College of Education, Sultan Qaboos University, Muscat, Sultanate of Oman.

Acknowledgements
AM carried out the literautr review studies, participated in the analysis part and drafted the manuscript. AMK carried out the statistical analysis. AH and FB designed the research instruments and validated them. All authors read and approved the final manuscript.

Acknowledgements
Funding was provided by Sultan Qaboos University (His Majesty Grant).

Competing interests
The authors declare that they have no competing interests.

Appendix

See Table 17.

Table 17 Attitude scale

No	Statement	Agree	Not sure	Not agree
1.	Software develop my self-learning skill			
2.	Software helps me to understand the Arabic language lessons			
3.	I feel bored when learning the Arabic language using the software			
4.	When learning through software, teacher does nothing			
5.	I respect a teacher who teach me Arabic language using the software			
6.	I can master reading skills when using the software			
7.	Software enables me to learn to read quickly			
8.	Teaching using the software makes the lesson fun			
9.	Use software increases my motivation/interaction with the lesson			
10.	Training on the use of software to learn Arabic causes me stress and fatigue			
11.	I feel a lot of enthusiasm when learning to read while using the software			
12.	Learning using the software leads to easier access to information and knowledge			
13.	Software help me gain speed to acquire difficult information			

Received: 7 April 2016 Accepted: 17 September 2016
Published online: 28 September 2016

References

Al Barakat A (2010) Effectiveness of a story strategy based training program on the development of Quranic assimilation of the basic schools' third-graders and their attitudes toward it. Umm Al Qura Univ J Educ Psychol Sci 2(1):391–451

Al Musawi A (2000) Technology effectiveness in improving teaching, learning, and communication skills: literature review and implications to Arabic education. Contemp Educ J 54:221–244

Al Musawi A, Al Hashmi A, Kazem A, Al Busaidi F, Al Khaifi S (2014) Perceptions of Arabic language teachers toward their use of technology at the Omani basic education schools. Education and Information Technologies, http://link.springer.com/article/10.1007/s10639-013-9305-5

Al Najjar M (1987) Treatment plan for reading and writing weakness in the fourth grade: experimental research. Educ Stud Number 4:53–67

Al-Busaidi F, Al-Hashmi A, Al Musawi A, Kazem A, Al-Khayfi S (2013) Standards for designing and producing linguistic software to address the weaknesses in reading at Omani first basic education cycle's schools, Paper presented at the international conference on languages, linguistics and society 2013 (ICLALIS 2013), Malaysia

Aldalalah M, Soon F, Ababneh W (2010) Effects of multimedia-based instructional designs for Arabic language learning among pupils of different achievement levels. Int J Hum Soc Sci 5(5):311–317

Al-Fori A (1999) Preparatory stage students' mastery of silent reading at Omani Muscat governorate, Unpublished MA, Sultan Qaboos University, Oman

Al-Hashmi J (2009) The effectiveness of computerized program in the treatment of reading aloud among the sixth grade students in the second cycle of basic education, Unpublished MA, Sultan Qaboos University, Oman

Al-Heela M, Ghoneim A (2002) The impact of regular and computerized educational games in dealing with reading difficulties among fourth grade students. Najah Univ J Res Humanit Stud 16(2):54–77

Al-Khatib H (2011) Technology Enhanced learning: virtual realities; concrete results. Case study on the impact of TEL on learning, European Journal of Open, Distance and E-Learning- EURODL, Issue 1, http://www.eurodl.org/?p=&sp=full&article=423

Al-Mawali H (2003) Evaluation of the Omani first secondary schools' students' mastery of reading aloud skills, Unpublished MA, Sultan Qaboos University, Oman

Almekhlafi AG, Almeqdadi FA (2010) Teachers' perceptions of technology integration in the united arab emirates school. Classrooms Educ Technol Soc 13(1):165–175

Al-Ramadhani S (1995) Common reading aloud mistakes among preparatory stage students of Omani Musandam governorate: its types, reasons, and solutions, Unpublished MA, Sultan Qaboos University, Oman

Al-Yazidi A (2006) The effectiveness of a computerized educational program to teaching Arabic reading on developing creative thinking skills of the fourth grade students in the Sultanate of Oman, Unpublished MA, Yarmouk University, Jordan

Awadh F (1999) The effectiveness of a proposed program on female students-teachers' skills' acquisition of reading weakness diagnostic therapeutic teaching skills for preparatory stage, Department of Curriculum and Instruction, Ain Shams University, Egypt

Awadh A (2010) The effectiveness of a proposed strategy in the treatment of reading, writing and achievement weakness in Arabic language for some basic education ninth grade students. Fac Educ J Univ Mansoura 73(1):320–374

Beauvois MH (1997) Computer-mediated communication: technology for improving speaking and writing. In: Bush MD (ed) Technology-enhanced language learning. National Textbook Company, Lincolnwood, pp 165–184
Blake RJ (1998) The role of technology in second language learning. In: Byrnes H (ed) Learning foreign and second languages: perspectives in research and scholarship. Modern Language Association, New York, pp 209–237
Cunningham A, Redmond M (2008) Instructional design and early language learning: cognition, creativity, and technology, American Association of Teachers of Spanish and Portuguese, pp 435–445
Dhir A, Alsumait A (2012) Examining the educational user interface, technology and pedagogy for Arabic Speaking children in Kuwait, paper presented at the 15th International Conference on Enterprise Information Systems (ICEIS) 2012
Ertem I (2010) The effect of electronic storybooks on struggling fourth-graders' reading comprehension. Turkish Online J Educ Technol TOJET 9(4):140–155
Goh Y, Ng A, Raja B, Wan A (2004) Technology and foreign language learning: student perceptions on the feasibility of using WBI (Web-Based Instruction) to supplement the on-campus foreign language courses in UiTM, Technical Report. Institute of research, development and commercialization, Universiti Teknologi MARA
Hoopingarner D (2009) Best practices in technology and language teaching. Lang Linguist Compass 3(1):222–235
Ismail S, Almekhlafi A, Almekhlafy M (2010) Teachers' perceptions of the use of technology in teaching languages in United Arab Emirates' schools. Int J Res Educ (IJRE) 27:37–56
Judson E (2010) Improving technology literacy: does it open doors to traditional content? Educ Tech Res Dev 58(3):271–284
Katz L, Carlisle J (2009) Teaching students with reading difficulties to be close readers: a feasibility study. Lang Speech Hear Serv Sch 40:325–340
Kim J, Samso J, Fitzgerald R, Hartry A (2010) A randomized experiment of a mixed-methods literacy intervention for struggling readers in grades 4–6: effects on word reading efficiency, reading comprehension and vocabulary, and oral reading fluency. Read Writ 23(9):1109–1129
Ministry of Education (2005) Reading weakness among first basic education stage: the phenomenon, reasons, and solutions, MOE, Muscat
OECD—Organization for Economic Co-operation and Development (2015) Students, computers and learning: making the connection. OECD Publishing, Pisa. doi:10.1787/9789264239555-en
Spreen C (2002) Preface: promises and pitfalls of technology in language learning. In: Spreen C (ed) New technologies and language learning: cases in the less commonly taught languages, National Foreign Language Resource Center. University of Hawaii Press, Honolulu, pp 7–20
To'eima R (2004) Language skills: their levels, teaching, and difficulties. Al-Fikr Al-Arabi Publishing, Cairo
Torgesen JK (2002) The prevention of reading difficulties. J Sch Psychol 40(1):7–26