TECHNICAL REPORT

Department of Computer Science

Investigation into the effect of Language on Performance in a Multimedia Food Studies Application

Trevor Barker, Sara Jones and Carol Britton

Technical Report No 324

Feb 1999

Investigation into the effect of Language on Performance in a Multimedia Food Studies Application.

Keywords: Multimedia, Pedagogy, Language skills, Multi-modal interface design

Authors: Trevor Barker, Sara Jones and Carol Britton
Department of Computer Science,
University of Hertfordshire, Hatfield AL10 9AB
David Messer
Department of Psychology,
University of Hertfordshire, Hatfield AL10 9AB

Author for correspondence:

Trevor Barker
Email: T.1.Barker@herts.ac.uk
Department of Computer Science
University of Hertfordshire,
Hatfield
AL10 9AB
01707 284 345

Abstract:

An investigation into how performance in a multimedia learning application was related to the level of language support available is reported. Teachers from a College of Further Education, National Vocation Qualification (NVQ) level 2 catering students and Higher National Diploma (HND) catering management students took part. After a language and subject pre-test, participants were randomly assigned to presentations of a multimedia catering course having either full, or no additional language support available. Immediately after completion, a post-test and two weeks later a re-test were taken.

No significant difference between staff or HND users with different levels of language support was found. Significant differences were found for the NVQ users with different language support. The results of this study are discussed in relation to language skills and the potential for the individual configuration of multimedia.

1.0 Introduction

Blank and Solomon [5] identified deficiencies in language skills in children and suggested that language deficiencies in children related to the lack of a system of symbolic thinking. It is likely that such serious deficiencies in language skills identified early in life would be carried on throughout the education process into adult life, and have a negative influence at each stage of development. Deficiencies in language skills might therefore lead to a poor educational experience and poor achievement in general.

This is especially true for students entering further education for several reasons. There is evidence that many learners have severe linguistic deficiencies when they enter college [2]. In many cases their first language is different from the language in which their selected course is delivered. Specialised vocabulary used in some courses may be a problem for many learners with technical and foreign terms and non-standard usage of words being common. There may also be problems of linguistic style, for example, the use of passive rather than active sentences and the use of past tense in technical writing.

It is sometimes assumed that using multimedia in learning with its use of image, video and animation, will compensate for deficiencies in learners' language skills. Petre [10], however, challenges this view and stresses the importance of text and spoken language in multimedia presentations. This leads to a high requirement for reading and listening skills in the learner. The multi-modal nature of the user interface places additional emphasis on the use of language in multimedia delivered learning. Complex written and spoken instructions are often involved in multimedia learning applications. McAteer and Shaw [7] recommend that authors pay particular attention to the use of language in multimedia applications and provide support when communication is in the form of written text in the application. Barron and Atkins [3] have found that listening skills are also important in multimedia learning applications. Poor use of language in the computer interface may also lead to usability problems in applications. Molich and Nielsen [9] stress the need for clear simple language in the design of computer interfaces.

The objective of the project reported here was to investigate the effect of language on performance in a multimedia learning application. The study was based on food studies courses taking place in a college of Further Education. Catering courses had been identified by language experts in the college as posing special language problems. There were many technical and foreign terms in regular usage and there was a high scientific component in some areas of food studies, for example, food hygiene and nutrition.

A network-delivered computer application was developed to deliver a course based on the food commodities. Teachers from a College of Further Education, National Vocation Qualification (NVQ) level 2 catering students

and Higher National Diploma (HND) catering management students followed a multimedia delivered learning application that provided different levels of language support. The aim of the investigation was to relate performance on pre-tests, post tests and re-tests to the levels of language support provided by the application.

2.0 Development of the software.

Three pieces of software were developed for this investigation.

1 A computer based tool to facilitate the creation of language differentiated materials (Language level tool).
2 A language testing tool to assess students' language skills
3 A catering commodities multimedia application able to provide differential language support for learners

2.1 Language level tool

A computer program was developed to determine the language levels of texts and narratives used in the course. This application was designed to measure the Adult Literacy and Basic Skills Unit (ALBSU) SMOG levels [11]. SMOG levels are used as a standard within many Further Education (FE) colleges to classify learners' reading, writing and listening skills. The SMOG level is calculated from the length of sentences and the number of syllables in words used in a text.

The language level tool was used to analyse sections of text and narrative used in the courses and to calculate the SMOG level of the language used. In this way it was possible to prepare text and narrative at a range of levels.

2.2 Language testing tool

The language test was based on a simple listening and gapping test as described by Vaughn [11]. The application tested simple listening and reading skills in the catering subject area.

The test developed followed closely existing language assessment and screening tests used routinely in the college based on the ALBSU SMOG test [11].

2.3 Multimedia application

The application was developed by a team of language, subject and computer specialists, using iterative prototyping and user centred methods. A description of the multimedia development process employed in the creation of the materials is given by Barker and colleagues [1].

The programme covered aspects of catering practical and theoretical work, including food science, food composition, food hygiene, storage, handling, nutrition, cookery and other related areas. The domain was selected to be relevant to as wide a range of students as possible. Catering specialists created a full specification for the application, which included text and narrative. After measuring the SMOG level of the language used, in the initial text and narrative, the application was differentiated to provide extra language support in the following ways.

1 Alternative words were provided. Instead of a long difficult word, shorter, simpler words were used.

2 Sentences were made shorter. Long sentences were cut up and presented as several smaller ones.

The effect was to create two versions of the text and narrative for the application, a high level version at ALBSU SMOG level 18 and a lower level version at level 14. These were used to create a prototype of the application differentiated at two language levels. Two pathways were provided through the prototype, a high level language route, set at SMOG level 18 and a lower level language route, set at SMOG level 14. The subject content covered in each pathway was identical.

In addition to reducing the SMOG level, additional language support was provided for the lower level pathway in the prototype by the following measures:

3 Sentences were made active rather than passive in the presentation with additional language support.

4 Additional hyper-linked glossaries and explanations were given where the language might be difficult, for example words like vitamin and protein were explained more in the extra support presentation

5 Additional images and videos were available in the extra supported presentation.

Only redundant information was provided by these additional measures to ensure information provided in both paths was identical.

The prototype was designed so that presentation could be varied within the application according to the values of variables held in individual configuration files for each user.

- Language level = 0 – High level presentation, (SMOG level 18) with no extra language support
- Language level = 1 – Low level presentation, (SMOG level 14) with extra language support provided

Sound presentation was configurable within the application in a range of ways to allow flexible use. For this investigation, sound was set on, repeatable and interruptible.

3.0 Method

Three groups of participants were involved in this experiment. Two groups of students, one following a Higher National Diploma (HND) and the other following a National Vocational Qualification (NVQ) level 2 in catering. A third group of non-catering lecturing staff also took part.

HND and NVQ students, despite their academic separation, follow a similar core course in food commodities as part of their normal curriculum. Differences in their academic levels however meant that these groups had different language support requirements.

The staff group had no vocational or subject experience of catering, yet would be expected to possess good language skills.

Table 1 below shows characteristics of the participants involved in the trial. The language score for each group is also shown in the table.

.Table 1

Language scores and groups of participants in the study

Group	N	Mean age	Age Range	Mean % Language test score
NVQ	32	17.9	16-23	60
HND	32	19.3	17-35	81
Staff	20	29.3	23-47	93

Each group was divided randomly into two equal parts. This enabled participants in each group to be assigned to either of the presentation regimes as follows:

Presentation a) without additional language support (none)

Presentation b) with additional language support (full)

Details of sub groupings and the language support presentations given are displayed in the table below.

Table 2

Additional language support, age and language test scores

Sub Group	Additional Language Support Given	N	Mean age	Age Range	Mean Language test score
NVQ a	None	16	18.3	16-23	62%
NVQ b	Full	16	17.4	16-22	57%
HND a	None	16	19.0	17-25	81%
HND b	Full	16	19.5	17-22	80%
Staff a	None	10	29.2	23-47	94%
Staff b	Full	10	29.8	23-45	92%

An analysis of variance was performed on the language test scores for all groups. There were significant differences ($p<0.01$) between the NVQ (a and b) groups and all other groups.

There was no significant difference between the NVQ a and NVQ b group ($p>0.05$).

4.0 Implementation

Participants were given a brief introductory talk prior to first use of the system. Immediately after induction, participants were administered the multimedia language test, followed immediately by the multimedia pre-test of 30 multiple choice questions.

Participants followed the course over a period of one week in open access computer areas. Students were supervised at all times by tutors who could provide additional help related to using the application rather than subject information.

Once the course had been completed, a multimedia post-test and user-evaluation was taken by all participants.

The user evaluation tool consisted of a set of 30 questions delivered on the computer in multimedia format. It measured how interesting the course material was, any areas of difficulty within the course and users' computer experience and familiarity with the use of multimedia hardware.

Two weeks later a supervised re-test was taken in multimedia format, delivered on a computer. The pre-test, post-test and re-test were in the same format and covered the same subject areas. Questions for these were selected randomly from a bank. A subject expert who assessed the tests during the software development process rated them to be of equal difficulty.

All results were saved securely and anonymously on a computer network. An extensive data log file was created for each subject throughout the course. This held information about navigation and time spent in each section of the course.

In summary, the implementation had the following stages:

- Initial language assessment test presentation
- Initial subject pre-test presentation,
- User configuration file created
- Course followed with prescribed language support
- Post-test presentation
- Evaluation of the application by users
- Re-test presentation two weeks after finishing the course.
- Data collected and analysed

5.0 Results

In this section, results obtained in the investigation and their statistical analysis is presented. Table 3 below presents the pre-test, post-test and re-test scores for groups following the commodities multimedia course. Results of the user evaluation questionnaire are also presented.

Table 3

Mean Pre-test, post-test and re-test and user evaluation scores for participants following the catering commodities multimedia course.

Group	N	Pre Test	Post Test	Re Test	Evaluation.
Possible score		(30)	(30)	(30)	(5)
NVQ a	16	11.13	14.44	12.5	3.25
NVQ b	16	12.13	18.56	14.38	3.68
HND a	16	15.81	21.44	18.25	3.10
HND b	16	16.06	19.94	17.69	3.32
Staff a	10	14.50	20.3	17.1	3.10
Staff b	10	16.60	22.6	18.5	3.20

5.1 Tests of assumptions

The statistical methods used in the data analysis assume that the observed covariance matrices of the dependent variables are equal across groups.

Box's Test of Equality of Covariance Matrices was employed to test the null hypothesis that they were equal. The observed value of $p > 0.05$ (0.85) compels us to accept the null hypothesis.

Mauchly's test was employed to test the sphericity of the data within groups which is also assumed within the methods used. The observed value of $p > 0.05$ (0.98) compels us to accept the null hypothesis and assume sphericity of data.

5.2 Analysis of Variance (ANOVA)

The means of the TRIALS variable (pre-test score, re-test score and post-test score) were subjected to an repeated measures ANOVA to test for the significance of any differences between them. Table 4 below shows the results of this analysis.

Table 4

Tests of Within Subject Effects

Results of repeated measures ANOVA performed on data from table 4.3

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
TRIALS	1087.727	2	543.864	74.74	0.000
TRIALS x GROUP	10.452	4	2.613	0.359	0.837
TRIALS x LANGUAGE LEVEL	32.177	2	16.088	2.211	0.113
TRIALS x GROUP x LANGUAGE LEVEL	14.818	4	3.704	0.729	0.729
Error (TRIALS)	1135.175	156			

Table 4 shows a significant difference between the means of the TRIALS variable ($p<0.001$). There were significant differences between the mean scores obtained in pre-test, post-test and re-test. Between subject effects were also investigated and results of this analysis are presented in table 5 below.

Table 5

Tests of Between Subject Effects

Results of repeated measures ANOVA performed on data from table 5.5

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	67519.376	1	67519.3	3010.4	0.000
GROUP	1135.176	2	567.588	25.306	0.000
LANGUAGE LEVEL	158.167	1	158.167	7.052	0.010
GROUP x LANGUAGE LEVEL	38.347	2	19.173	0.855	0.429
Error	1749.438	78	22.429		

Table 5 shows significant between subject effects for the GROUP and LANGUAGE LEVEL variables.

Post hoc comparisons were performed using Tukey's test in order to investigate differences in the group means. Results of the post hoc comparisons are shown in table 6 below. Significant differences ($p<0.05$) are indicated by the presence of a '*' in the column.

Table 6

Tukey's Test

Post hoc pair-wise differences in the means for the data presented in table 3

'*' in a column indicates a significant difference at $p<0.05$

Trial	Group	Language support Condition None = no support Full = full support	1	2	3	4	5	6
Pre Test			NVQ a	NVQ b	HND a	HND b	Staffa	Staffb
	NVQa	None			*	*		*
	NVQb	Full			*	*		*
	HNDa	None	*	*				
	HNDb	Full	*	*				
	Staffa	None						
	Staffb	Full	*	*				
Post Test								
	NVQa	None		*	*	*	*	*
	NVQb	Full	*					
	HNDa	None	*					
	HNDb	Full	*					
	Staffa	None	*					
	Staffb	Full	*					
Re Test								
	NVQa	None			*	*	*	*
	NVQb	Full				*		*
	HNDa	None	*					
	HNDb	Full	*	*				
	Staffa	None	*					
	Staffb	Full	*	*				

6.0 Discussion

Tukey's test indicated that there were significant differences between the performance of individual groups on pre-test, re-test and post-tests under conditions of different language support. Differences in pre-test scores between the NVQ and HND / staff groups ($p<0.05$) were likely to be due to different abilities, experience and prior knowledge between the groups. In fact HND and staff groups performed better on average in all tests than NVQ groups. This was significant ($p<0.05$) in all cases except for the staff a group in the pre-test. Staff groups performed no differently from HND groups on average in all tests under both language conditions.

A significant difference was found in the re-test results between NVQ groups under the two conditions ($p<0.05$). The NVQ group receiving additional support and lower level language presentation performed significantly better on the re-test than the NVQ group at the higher language level with no support. The provision of additional language support therefore, was most effective for NVQ learners who scored lowest of all groups on the

language test. Differences in post-test scores between the two NVQ groups therefore, were ascribable to the additional language support provided within the application.

There are implications of these results for the configuration of multimedia learning materials. When learners have high level language skills, the provision of additional language support is not likely to be effective in improving performance on a multimedia course. When learners have language deficiencies, then it is of benefit to learners to provide additional support and to present language at the appropriate level.

All participants undertook an evaluation of the package in the form of a multimedia presented questionnaire. Table 4 shows that groups with the supported presentation on average scored the package higher than those with the unsupported presentation. This difference was greatest between the NVQ groups, those with additional support scoring it higher than those without the benefit of this. This difference however was not significant ($p>0.05$). Although there was also no significant difference between the evaluation scores ($p>0.05$), HND and staff groups with additional language support evaluated the application lower than those without additional support. The slightly lower evaluation scores for the higher level groups following the lower level language pathway providing additional support suggests that this may cause some level of de-motivation of learners with good language skills, though there was no statistical support for this idea.

Failure to find significant differences in the perceived quality of the application in three diverse groups of learners suggests that the provision of differential language support in an application is an important way to tailor it to a specific group of users. In this way learners with poor language skills are likely to benefit from the additional language help available and perform better. Learners with good language skills, who were shown not to benefit from additional language support, may prefer a language presentation at their ability level.

The use of sound has been shown to add to a learning presentation, not only in terms of content and information presentation effects, but it may add audio cues and interest to the application as with music and sound effects, McAteer and Shaw [7]. The use of sound has been suggested to offer benefits when language skills are poor. Barton and Dwyer [4] report that subjects with high verbal skills do not benefit from the addition of audio information in learning applications. They do suggest however, that subjects with lower verbal skills might benefit from textual / audio redundancy in learning. Kenworth [6] supports this view, suggesting that poor readers benefit from hearing text presented.

Meskill [8], suggests that the control of the rate of language presentation in multimedia applications allows the retention of language chunks in short term memory. This could in itself be important in improvement in performance in learning when listening skills are limiting. Meskill emphasized the potential of multimedia in language learning and sees listening as a skill integral to overall communicative competence.

Co-ordinated visual, aural and textual information employed in multimedia can provide clues to the meaning of the written and aural text according to Meskill. The results of the investigation reported here indicate that the presentation of information at the appropriate language level assists in this process.

The investigation showed that performance in a multimedia learning application is improved for learners when it is presented at the most appropriate language level. It is also suggested that failure to do this will result in less than optimum performance, either by de-motivating learners with good language skills or by setting the language level too high to be understood. The individual configuration of multimedia presentations is an important area for future research.

7.0 References

1. Barker, T., Jones, S., Britton, C. and Messer, D., J. (1997) Creating Multimedia Learning Applications in a Further Education Environment, Technical Report No. 271, University of Hertfordshire, Division of Computer Science, January 1997

2. Barnsley, M. (1996) Language screening test results, Language Centre, Internal report. , Waltham Forest College, London, E17 4JB

3. Barron, A. and Atkins, D. (1994) Audio instruction in multimedia education: Is textual redundancy important?, Journal of Educational Multimedia and Hypermedia, 3, 295-306.

4. Barton, E., A. and Dwyer, F., M. (1987) The effect of audio redundancy on the students' ability to profit from printed-verbal visualized instruction, International Journal of Instructional Media, 14, 93-98.

5. Blank and Solomon, (1969) How shall the disadvantaged be taught? Language in Education, Open University Press.

6. Kenworth, N., W. (1993) When Johnny Can't Read: Multimedia Design Strategies to Accommodate Poor Readers, Journal of Instructional Delivery Systems, Winter 1993, 27-30

7. McAteer, E. and Shaw, R. (1995) The Design of Multimedia Learning Programs. Establishing Multimedia Authoring Skills in Higher Education. EMASHE Group publication, University of Glasgow.

8. Meskill, C. (1996) Listening skills development through multimedia, Journal of Educational Multimedia and Hypermedia, Vol 5, No. 2

9. Molich, R. and Nielsen, J. (1990) Improving a human-computer dialogue. Communications of the ACM, 33(3), 338-348.

10. Petre, M. (1995) Why looking isn't always seeing. Readership skills and graphical programming. Communications of the ACM 38 (6) 33-44

11. Vaughn, J. (1995) Assessing Reading: Using cloze procedure to assess reading skills, Adult Literacy & Basic Skills Unit, London.