Lexis that rings a bell:
on the influence of auditory support in vocabulary acquisition

This empirical study investigated the effectiveness of auditory support in vocabulary learning by comparing acquisition and retention of lexical items studied using a traditional paired-associates memorisation technique to results achieved using an audio-supported paired-associates technique. The subjects were 88 Korean university students. Results indicated that the audio-supported vocabulary learning approach leads to significantly higher rates of acquisition. This advantage was retained over the two months following treatment. The largest difference was noticed in pronunciation. Furthermore, it was found that success in the audio-supported approach was not significantly dependent on learning preferences and that the approach enjoyed a higher level of acceptance among subjects than non-auditory paired-associates memorisation.
key words: auditory vocabulary acquisition Korean English

Introduction

Background and Motivation

The great importance of lexis to second language (L2) proficiency is reflected by the large number of recent studies investigating aspects of L2 vocabulary acquisition. A host of different acquisition techniques, both incidental and explicit, and a range of aspects influencing acquisition have been investigated. Studies investigating explicit techniques have, for example, looked at paired-associates memorisation (usually in contrast to other methods as in Lotto and de Groot 1998, Hermann 2003, Sagarra and Alba 2006, de Groot 2006), classroom vocabulary activities (Cho 2002, Rott, Williams and Cameron 2002, de la Fuente 2006) or the keyword method (Bird and Jacobs 1999, Shapiro and Waters 2005). Studies of incidental vocabulary acquisition have looked at acquisition through reading (Pitts, White and Krashen 1989, Paribakht and Wesche 1996, Waring and Takaki 2003 among others) and, to a lesser degree, through listening (Ulanoff and Pucci 1999, Vidal 2003). Effects of background music, similarity to L1, concreteness, frequency (de Groot 2006) or context (Webb 2007) have also been studied.

No scholarly attention, however, has been paid to explicit vocabulary learning techniques employing auditory support such as the 'listen to tape of word lists' strategy listed in Schmitt's taxonomy of 58 vocabulary acquisition strategies (1997: 208). This is somewhat surprising in view of the increased interest in the auditory medium brought about by the popularity of highly portable digital audio players and of podcasting. The latter was discovered as a tool for language learning early on, and its current use as a

vocabulary learning tool is exemplified by such podcasts as Just Vocabulary (2007), Français Interactif: liste de vocabulaire (2007) and Vocabulario Español (2007) among many others. Further, conventional commercial products for L2 vocabulary learners such as the VocabuLearn series (for example Anon., 1987) on CD, tapes or MP3 and new commercial vocabulary learning software, such as the Clip2go series of digital audio player software (Clip2go Basic Vocabulary, 2006) employ auditory approaches. The present study seeks to address the imbalance between the evident interest in explicit audio-supported vocabulary learning on the one hand and work investigating the effectiveness of such approaches on the other.

Theoretical Context

Lexical Knowledge

Knowledge stored in the mental lexicon is multi-faceted. Entries can encompass phonological and orthographic forms, links to relevant semantic information (including polysemy, homonymy, hyponymy, antonymy and other meaning relations), grammatical and morphological properties as well as information on collocation, style, connotation and register (Nation 1990: 30-33, Laufer 1997: 141 among others). While no evidence has emerged for there being a particular order of acquisition such as is now accepted regarding the acquisition of grammar (Ellis 1994: 21), learners inevitably move from knowledge of one or two aspects toward knowledge of most aspects and from basic knowledge of an aspect to more extensive knowledge of it. The term 'semantization process' has been suggested to draw attention to the various simultaneous processes involved in acquiring vocabulary knowledge (Henriksen 1999: 307). Extensive knowledge of all aspects of a word will not be required for all items in the lexicon and words are often learned partially (Laufer 1997: 142; Waring and Takaki 2003: 133). Further, there is some evidence for differences in accessibility, i.e. between receptive and productive knowledge, with the former ordinarily thought to precede the latter (Melka 1997: 84). It follows from the multi-faceted nature of vocabulary knowledge that acquisition methods ought to involve a range of aspects and that to credibly establish effects of different acquisition approaches, several aspects will need to be measured (so also Waring and Takaki 2003: 133, Wesche and Paribakht 1996 and others).

Input Channel

An audio-supported approach to explicit vocabulary learning differs from similar explicit approaches essentially in the channels in which the input is provided. While available psycholinguistic experiments have not explicitly addressed the effects of differences in channel in an L2 acquisition context, there is evidence that the addition of auditory input in an L2 vocabulary learning situation may result in more effective processing and transfer of input to long-term memory:

In Baddeley and Hitch's influential model of working memory (Baddeley and Hitch 1974; Baddeley 2003, 2007), verbal short-term memory (the phonological loop), where language comprehension and processing takes place, is comprised of a phonological store and an articulatory rehearsal mechanism. The phonological store was found to hold information only for seconds (Baddeley 2003: 191). To retain form longer and pass it on to long-term memory, traces in the phonological store need to be refreshed by the articulatory rehearsal mechanism through subvocal repetition

(Baddeley 2003: 191). That explicit L2 vocabulary learning in particular relies on the articulatory rehearsal mechanism was shown, for example, by Papagno et al. (1991) who found that a disturbance of the rehearsal mechanism significantly impacted on the ability to memorise L1 – L2 word pairs. Work by Service suggested that 'the ability to represent unfamiliar phonological material in working memory underlies the acquisition of new vocabulary items in foreign-language learning' whether input is auditory or visual (1992: 21). Baddeley found that auditory input gains automatic access to the phonological store (2007: 9). Visual, such as orthographic, input, however, is first received into the visuospatial sketchpad (the visual short-term memory) where an image-based representation of the information is kept. To access the phonological loop, such information first has to be subvocalised (Baddeley 2003: 191). Figure 1 is a simplified illustration of this process based on Baddeley (2003: 193).

-> Figure 1 about here

It follows that if there are difficulties in obtaining the phonological form of an orthographic representation (as might well be the case with unfamiliar L2 lexis provided exclusively through the visual channel), this can be expected to have a negative impact on rehearsal and ultimately transfer to long-term memory. We therefore expected overall scores of the audio-supported technique to be minimally on par with the non-auditory control treatment, possibly somewhat higher. On the other hand, no clear evidence is available on whether one of the channels would favour certain aspects of vocabulary knowledge: it has yet to be established whether the auditory channel is most useful in the acquisition of the phonological form and the visual channel of the orthographic form, as common sense might lead one to assume.

Learning Preferences

Since the 1970s, the idea that different people have different ways of going about learning and different ways in which they learn best, often referred to as 'learning style' or 'learning preference', has had an increasing influence on educators. No single taxonomy of learning styles exists, but the distinction between field-dependent and field-independent learners (derived from Witkin et al. 1971, 1977) and categories such as visual, auditory and kinaesthetic learning preferences (derived from Neuro-Linguistic Programming (NLP), esp. Bandler and Grinder 1979 and Grinder 1991) are now commonly referred to in the language teaching literature. While audio-supported vocabulary learning would appear to suit learners with a preference for auditory learning (and more conventional visual approaches might favour visual learners), research has not so far been conducted on whether this is indeed the case and consequently whether audio-supported vocabulary learning could only be recommended to auditory learners. Adopting a conceptualisation of learner preference based on NLP, the present study seeks to shed light on this issue as well.

Research Questions

Based on the above considerations, this study sought to address the following research questions:

1: In terms of acquisition and retention, how does an audio-supported approach to explicit vocabulary learning compare to a similar non-auditory approach?

2: Is auditory support particularly useful in aiding the acquisition of the phonological form of lexical items and relatively less useful in aiding the acquisition of the orthographic form, as might be expected a priori?

3: Does the degree of success using the audio-supported method correlate with learner type? (i.e. Do learners with a preference for auditory learning benefit more from this method than others?)

The study further hoped to provide indicators of the level of acceptance of the audio-supported approach among learners. For question 1, the audio-supported approach was expected to yield results minimally on par with the non-auditory control approach. For the remaining questions, previous research did not supply sufficient basis for formal expectations.

Method and Procedure

Overview of Basic Design

To answer research questions one and two, subjects were asked to memorise 20 previously unknown words during four time-limited treatment sessions. Ten words were studied using an audio-supported approach and ten words using a non-auditory one, as described below. Three post-tests were administered after treatment: a 48-hour post-test (t1), a ten-day post-test (t2) and a two-month post-test (t3). Gains in lexical knowledge under each condition were measured and compared. For research question three, the results of a learner style questionnaire were compared to participants' scores in post-tests. Data on the level of acceptance of the audio-supported approach among participants was collected using a questionnaire.

Participants

The participants in the study consisted of three classes of first year university students (typical age: 19) following a core EAP course at a Korean university (n = 88). Most students were at an intermediate level of English (level B1 of the Common European Framework [Council of Europe, 2001]) as established by the researcher after teaching the participants for one semester, with a few students at upper intermediate (B2) and pre-intermediate (A2) levels. All participants shared the same L1 background (Korean). Due to absences and other reasons described below, the scores of only 62 participants could finally be used in the study.

Development of Instruments

Treatment Design

The 20 previously unknown words used in treatment sessions were low-frequency words and two academic lexical items. Initially, a list of 30 items was compiled: 22 words from the latest version of Paul Nations' Vocabulary Level Test (VLT) in its Korean bilingual version (Nation 2007, personal communication), 20 at the 10,000 word level (i.e. words from among the 10,000 most frequent words of English, excluding the 5,000 most frequent words) and 2 at the academic level. Eight low-frequency words (not among the 5,000 most frequent words of English) were added by the researcher. Care was taken to avoid cognates or false friends and no

morphologically complex items or items longer than three syllables were admitted (except for the morphologically complex but semantically non-transparent 'blue-blooded'). In a second step, L1 glosses (translations, or, where impossible, short explanations) were added making use of the VLT's existing glosses and a bilingual dictionary for the remaining words. Glosses were then checked for accuracy and clarity by a native Korean speaker. Next, a pre-test was created containing only the glosses, with a space for the target word. This 30-item pre-test was administered to all participants who were asked to supply the English words which they thought fitted the glosses. The participants were asked to write down up to three English words if in any doubt as to which particular one was intended. As with all written instructions to participants in this study, the instructions were given in their L1 to maximise clarity.

Based on pre-test results, 20 unknown vocabulary items were chosen for treatment. Misspelled but recognisable target words were considered known. Ten students displayed fewer than 20 unknown words and their subsequent results were therefore excluded from the study. The 20 unknown items were divided into two sets of ten (set 1 and set 2). In order to achieve the maximum possible equality between the sets with regard to difficulty of acquisition, the factors of word-length (three one-syllable words in each set, no words over three syllables), difficulty of phonological form (such as occurrence of phonemes that typically cause difficulties for L1 speakers of Korean, with reference to Lee, 2001, p. 325-6) and difficult orthography (here the researcher's own judgement was relied upon initially and later adjusted in light of the error analysis described below) were controlled in each set in addition to the criteria applied at the initial selection (that is frequency, morphological simplicity, number of syllables and non-cognateness). This was further refined by an analysis of errors by item in a pilot study (n = 7, near-identical procedure as main study) which additionally allowed for other, less well-explored, difficulty factors to be taken into account. After minor adjustment to the assignment of words to each set, sets one and two arrived at a closely similar number of total errors of each type as measured in the pilot. To preclude the possibility of a remaining bias based on set, one group of participants (n=27) studied set 1 in auditory mode and set 2 in the non-auditory control mode whereas the other group (n=35) did the reverse. Table 1 illustrates the combinations of group, mode and set used.

-> table 1 about here

Under auditory mode, participants were supplied with the phonological form of the target word and two example sentences via audio. This information was supplied visually on a handout under the non-auditory mode. The orthographic form and an L1 translation or explanation were supplied visually in both auditory and non-auditory mode, as detailed in table 2. Appendix A contains treatment handouts showing the items of sets 1 and 2, example sentences and phonemic transcriptions used. Example sentences for all items were taken (and in some cases adapted) from the Cambridge Advanced Learner's Dictionary (Gillard [senior commissioning ed], 2003), L1 glosses were the same as on the pre-test. It should be noted at this point that the set-up described here did not test an exclusively auditory vs. an entirely visual learning technique but rather the contribution of the auditory element to a non-auditory technique (i.e. auditory support). While an investigation of exclusively auditory learning might have been of interest from a purely theoretical point of view, in practice it is

difficult to imagine that a purely auditory approach, without recourse to visual input of at least the orthographic form, would be used by many learners. Concurrently, it is worth pointing out that in accordance with the design specifications, each type of information (as listed in table 2) was supplied through only one channel and thus no doubling of information occurred.

-> table 2 about here

Post-Test Design

Due to the multi-faceted nature of lexical knowledge, accurately measuring it is a non-trivial challenge. Some approaches being used rely on self-reporting (The Eurocentre's Vocabulary Size Test: Meara and Jones, 1990), others on multiple choice questions (most notably the VLT, Nation, 1990 and subsequent revisions) and yet others require participants to demonstrate knowledge of target words by supplying written data such as a definition (Cho, 2002; Waring and Takaki, 2003; de Groot, 2006, etc.) or an example sentence. To overcome the oversimplification of a binary distinction (known/unknown), scalar measures have been devised such as the widely-known Vocabulary Knowledge Scale (VKS) (Paribakht and Wesche, 1993, 1996).

Though still suffering from some of the limitations common to all vocabulary assessment methods, the assessment scheme used for the present study sought to improve on the VKS in key areas while retaining a similar format. All three post-tests followed identical format and assessed word knowledge through measuring six aspects which scored one point each (thus a maximum of six points could be scored per word):

1) Word recall (**w**): ability to produce a written, recognisable form of the target word from an L1 prompt of one of its meanings. The form provided did not need to be the correct orthographic form; the prompts were identical to those used on the pre-test and treatment handouts. The marking was deliberately kept generous.
2) Spelling (**sp**): ability to produce the correct orthographic form of a successfully recalled word
3) Grammatically correct use in a sentence (**gr**): ability to use the recalled target word in a sentence with absence of grammatical errors. Only errors pertaining to the target word were considered and the problems of missing third person –s, article usage and orthography, even when in the vicinity of the target word, were not marked as errors because of their likely systematic nature. Use of the target word in a different part of speech from the one used in treatment resulted in zero points awarded because any knowledge so demonstrated would not have been the direct result of treatment. The lowest score was zero even if more than one target word-related error was found or no sentence was provided.
4) Semantically and collocationally appropriate use of the recalled target word in a sentence (**sm**): absence of target word-related semantic or collocational oddity in the sentence provided. Sentences did not have to demonstrate knowledge of the meaning of the target word but contradictory or nonsensical example sentences or those where the target word was used in a different part of speech were awarded zero points. The final two aspects tested phonological form:
5) Stress (**st**): here one point was awarded for correct stress pattern.
6) Phonemes (**phon**): a further point was awarded for absence of phonemic error.

Criteria *sp*, *gr* and *sm* depended on a successful word recall (*w*) from the L1 prompt (though not on correct spelling), thus measuring productive knowledge. This condition served to desensitise the measure at the lower end, setting the scoring threshold relatively high. It also allowed for the testing of the orthographic form while avoiding the validity and reliability problems of trying to measure knowledge of meaning via synonyms or L1 translations which possibly require a different skill set. While successful word recall is a plausible pre-requisite to correct orthographic form (*sp*) and for its use in a sentence (*gr* and *sm*), no other assumptions of linearity were made (i.e. the criteria are otherwise independent of each other in that a participant scoring under the *w*, *gr* and *phon* criteria but not under *sp*, *sm* and *st* represents no problem to the system just like a participant scoring only under *st* and *phon*, thus demonstrating the facet of word knowledge of being able to produce the correct phonological form when prompted with the orthographic form). Hence knowledge of the target word in any of the measured categories was counted without the imposition of an order of acquisition inherent in a linear model.

The written parts of post-tests (i.e. criteria *w, sp, gr, sm*) were initially scored by either the researcher himself or a research assistant. Tests scored by the research assistant were subsequently checked by the researcher for consistency. After several months, all tests were re-marked by the researcher in one sitting to ensure consistency in the application of marking principles. The audio parts of post-tests (the criteria *st* and *phon*) were scored twice by the researcher while listening to a recording (for procedure of recording see below). The pilot showed that when reading the list of target words, subjects often used 'upspeak' (i.e. raising their voice toward the end of a word) which made stress placement identification somewhat difficult. For the main study, therefore, full stops were placed at the end of each target word on the list which largely eliminated the problem. In the few remaining instances of upspeak, stress placement was marked wrong if any doubt existed. Phonemic errors were marked as such whenever individual phonemes were mispronounced, whether or not actual or potential confusion existed.

Questionnaires

For research question three, a learning preferences questionnaire was administered to all participants at the end of the first post-test. Wingate's (2000) questionnaire, slightly abbreviated and translated into participants' L1, was used to position respondents in relation to their preferences of auditory, visual and kinaesthetic learning styles. To measure levels of acceptance, an attitude questionnaire in participants' L1 was administered immediately following the last treatment session (English version shown in table 7 below).

Procedure

An overview of the procedure is given in figure 2.

-> Figure 2 about here

Treatment

After the pre-test (described above), four treatment sessions were administered in the space of eight days. Treatment sessions were administered during timetabled lectures under the supervision of the researcher. During treatment sessions, participants were asked to memorise the two sets of ten words as described above. Treatment handouts

(reproduced in appendix A) were designed in such a way that information for set 1 was printed on one side and set 2 on the other, preventing participants from looking at sets not currently being studied. The handouts were collected after each treatment session. Seven minutes of each treatment session were spent in auditory mode (which allowed for going through the set of ten items three times) and seven minutes in non-auditory mode, bringing the total length of each treatment session to fourteen minutes. Participants were told that while studying in auditory mode, they were free to either look at the handout or not as they saw fit. In auditory mode, all participants heard the recorded information which was audible throughout the room at speaking volume. If a participant had ignored the audio, which did not appear to be the case, this would have reduced their post-test scores under auditory mode because they would have missed out on example sentences and information on pronunciation. At the beginning of each treatment session, participants were additionally instructed not to write anything down and not to study target words outside of class. It was also made clear that the exercise was in no way part of the assessment for their course.

Post-Tests

Three post-tests were administered under the supervision of the researcher and participants were informed that they could opt out of taking the tests if they so chose at no disadvantage. Post-test 1 (t1) was administered 48 hours after the last treatment session, post-test 2 (t2) 10 days after the last treatment session and post-test 3 (t3) after two months.

Post-tests were administered in two phases. Phase one consisted of the completion of a written test (appendix B) which supplied the L1 prompt (identical to those used during treatment) and space for the target word followed by a longer space for the example sentence. The test order of words was such that a set 1 word was always followed by a set 2 word but the order within sets was randomised (i.e. different from the treatment order). The time limit for the written test was set at 15 minutes, which appeared sufficient. Any shortage of time did not affect the sets unequally due to the alternation of words from each set. Participants were asked not to supply example sentences identical to those used during treatment. Phase two, immediately following phase one, consisted of each participant being separately called out of the lecture hall and asked to read out loud a list of the 20 target words. This was recorded on tape. The time-consuming nature of phase-2-testing, coupled with the facts that subjects were only available during time-tabled lessons and that post-tests had to be administered at strictly time-limited intervals meant that phase two could only be administered for a random subset of participants and only at t1 and t3.

Participant Numbers

The scores of three participants were excluded as outliers because on t1 they did not manage to recall a single word studied in non-auditory mode. This was taken as indicative of a lack of participation in treatment sessions and their inclusion would therefore have skewed results. Furthermore, only scores of participants who were present at all post-tests (t1, t2, t3) were considered. Individual participants' identities were tracked using the last three digits of their university-assigned student ID numbers. In instances where participants opted out of providing ID and no identification could be made on the basis of handwriting, the scores were excluded from the study. Together with the scores of ten participants excluded on the grounds of pre-test results, this

brought the number of participants considered to 62 for written test parts and 45 for pronunciation test parts.

Results and Discussion

An analysis of scores by set showed that there was no significant difference between set 1 and set 2 on any of the three post-tests ($p > .05$ on a two-tailed t-test). This held true both overall and in each of the assessment criteria configurations (i.e. *w; w+sp; w+gr; w+sm; w+sp+gr+sm; st; phon; st+phon; w+sp+gr+sm+st+phon;* here as elsewhere, criteria *sp*, *gr* and *sm* are not analysed in isolation because of their dependence on *w*). Table 3 shows total scores and mean scores by set for t1. Thus, 'set' was not a variable which needed to be traced and participants' auditory scores could be compared with their non-auditory scores regardless of which set they studied in which mode. Participants thus served as their own controls and any differences between groups of participants were inconsequential.

-> Table 3 about here

Acquisition

Figure 3 (below) shows the scores on pre and post-tests for *w+sp*.

-> Figure 3 about here

Results of a comparison of mean scores in auditory and non-auditory mode are shown in table 4. Two-tailed t-tests on auditory and non-auditory scores revealed significantly superior scores for auditory mode over the non-auditory mode across the three tests and across criteria (an alpha-level of .05 is used throughout this study). Following Cohen's (1988) glosses, t1 showed small to medium effect sizes (d between 0.41 and 0.57) for all criteria configurations, except pronunciation: *st+phon* displayed an overwhelmingly large effect of d = 1.21. Effect sizes grew with time from a medium size of d = 0.57 on t1 (overall), to d = 0.70 on t2 (*w+sp+gr+sm*) and d = 0.78 on t3 (overall). The difference between auditory and non-auditory scores thus became more pronounced over time right across the different criteria. Interestingly, the pronunciation criteria were alone in seeing their effect sizes diminish somewhat over time (*st+phon*: d = 1.21 on t1 to d = 0.95 on t3) but together remained the criteria with the largest effect. Thus our expectation of auditory mode scores being minimally on par, but possibly higher than the scores under the non-auditory condition was confirmed. While we suggest this is due to the nature of verbal processing which favours techniques with auditory support, further research is needed before this can be confirmed.

-> Table 4 about here

Looking at the scores for auditory and non-auditory conditions by participant (table 5), it is clear that higher auditory scores were not only achieved overall but also by the majority of individual participants. Again, percentages of participants scoring higher under the auditory condition increased over time and the largest percentages were recorded on t3 where overall 84.4% of participants scored higher overall under the auditory condition.

Retention

The differences between t1 and t3 (table 4) show the rate of forgetting which, while larger under non-auditory than auditory mode (except for pronunciation), was not significantly different between conditions (p > .05). Under all criteria but *st* and *phon*, as well as overall, significant forgetting had occurred between t1 and t3 (as well as between t2 and t3, though not between t1 and t2) under both auditory and non-auditory conditions. Figure 4 shows scores over the three (two) tests for *w*+sp+gr+*sm, st+phon* and overall (*w+sp+gr+sm+st+phon*). This indicates that mode (auditory or non-auditory), while affecting acquisition, did not significantly influence long-term retention rates.

Pronunciation and Spelling

As expected, the pronunciation criteria benefited most from auditory support. *St+phon* showed consistently the largest effect sizes of all criteria (t1: d = 1.21, t3: d = .95). Although the effect of the auditory mode diminished somewhat over time, scores for the pronunciation criteria, contrary to the behaviour of the other criteria, saw no significant changes between t1 and t3 either in auditory nor non-auditory mode with the result that the auditory mode maintained its large advantage over the non-auditory under *st* and *phon*. The analysis by participants (table 5) shows that while the proportion of participants scoring higher on *st+phon* in auditory mode dropped from 80% in t1 to 66.7% in t3, it is still firmly above 50%. Furthermore, the percentage of participants scoring lower under the auditory condition was the lowest of all criterion configurations on t3 (*st+phon*: 13.3%).

Spelling, measured in conjunction with word recall (*w+sp*), while showing a significant advantage of the auditory mode over the non-auditory mode, displays the lowest d values of all criteria combinations on t1 and t2 and a relatively low value also on t3. Under the analysis by participants (table 5), adding *sp* to *w* had a polarising effect: the number of participants who scored higher under the auditory condition rose in all tests (56.5% to 58.1% on t1, 62.9% to 67.7% on t2 and 69.4% to 74.2% on t3), but the number of participants scoring lower also rose, leading to a fall in the number of participants who scored the same or higher under the auditory condition. Thus we can say that while the *sp* criterion may be a relative weakness of the audio-supported approach, it does not detract from the overall superiority of results under the auditory condition. Isolating the *sp* criterion (i.e. looking at the number of correctly spelt words as a percentage of correctly recalled words; *sp/w*) was only possible for t1 since on t2 and t3 there was a significant difference between sets and groups for *sp*/w). Table 6 shows the *sp* criterion in isolation for t1 where we find the non-auditory condition with an insignificant advantage. The comparatively high standard deviation for auditory mode shown in table 6 is caused by a participant who did not spell any recalled words correctly. No such scores of 0 occurred under the non-auditory condition. If the score is removed, and significance calculated again on this basis, the *p* value is higher yet,

indicating that the insignificance was not a result of the high standard deviation of the auditory scores.

-> *Table 6 about here*

Since the orthographic form was actually supplied visually in both auditory and non-auditory modes, this result is not a surprise. We can conclude that as long as the orthographic form is supplied visually, the audio-supported approach does not perform significantly worse under the criterion of spelling.

Level of Acceptance

Table 7 below shows the results of the attitude questionnaire. 72.9% of respondents deemed the audio-supported method more effective while 54% said that, if given a choice, they would prefer to use the audio-supported method rather than the non-auditory method in the future. Further, respondents also thought that the audio-supported method was less work and more fun than the non-auditory one. Thus it may be asserted that the audio-supported method enjoyed a somewhat higher level of acceptance among participants than the popular paired-associates memorisation technique as exemplified by the non-auditory method.

-> *Table 7 about here*

Learning Preferences Correlations

A comparison of participants' scores in the learner preference questionnaire with their test scores displayed the highest correlations when learner preference scores were compared to the difference between auditory test scores and non-auditory test scores. Still, significant correlations were found only on t1, where a medium positive correlation between test score differences and auditory learning preference scores were found overall (r = 0.332, *p* = .026). The scatter diagram of figure 5 illustrates this relationship. The difference between modes on word recall and spelling (*w+sp*) displayed a small and only just significant positive correlation with auditory learning scores on t1 (r = .261, *p* = .040). Statistically insignificant, small positive correlations were detected between auditory learning preference scores and test scores of most criteria and tests. Small and insignificant negative correlations were found between visual learning preference scores and test score differences on most criteria and tests. Kinaesthetic learning preference scores on the whole did not seem to correlate at all with post-test scores. Table 8 gives an overview of correlations and their significance levels.

-> Table 8 about here
-> Figure 5 about here

The data does not suggest a strong correlation between reported auditory learning preference and score differences (auditory minus non-auditory scores) on post-tests. Learners with auditory preferences cannot therefore be said to benefit significantly more from the audio-supported approach than others, especially in the longer term. A larger sample size might be able to find a higher number of significant correlations, although it is unlikely that they will prove to be strong. Precluding for the moment the

possibility that the questionnaire used to place learners in relation to their learning preferences was inadequate, it seems that success in the auditory method does not significantly depend on (self-reported) learning preferences, especially after some forgetting has taken place. A possible reason could lie in the phonological nature of vocabulary processing, discussed above, that holds for all learners and hence learners with auditory preferences would not be expected to perform significantly better. Future work will need to determine more clearly whether that is indeed the case. For now, the audio-supported approach appears recommendable to all learners regardless of learner preference.

Summary of Results and Conclusions

The audio-supported approach used in this study has led to a significantly greater increase in the vocabulary knowledge of participants than was achieved using the non-auditory paired-associates memorisation method. The advantage of the audio-supported approach was maintained over the period of two months that followed treatment. As expected, the auditory support was found to be especially useful for the acquisition of the phonological form, where the difference from the non-auditory condition was the largest ($d = 1.21$ on t1 and $d = 0.95$ on t3). While the criterion of orthographic form showed insignificantly lower results under the auditory than the non-auditory condition on t1, this did not detract from the overall superiority of the audio-supported approach.

It was also found that success in the audio-supported approach did not correlate strongly with any learning preferences as established through the questionnaire used in the study, though a small positive correlation of auditory learning preference with the difference between auditory and non-auditory scores was found on t1. We suggest that this is due to the nature of verbal processing favouring auditory input, overriding learning preferences.

The level of acceptance of the auditory method as reported by participants was superior to non-auditory paired-associates memorisation.

Implications and Limitations

In view of the results detailed above, vocabulary acquisition methods with auditory support similar to the one used in this study can be recommended to learners as an excellent method, with an advantage over the popular and effective paired-associates memorisation technique. This pertains to situations where explicit teaching/learning approaches are appropriate. Since evidence has emerged to suggest an audio-supported approach is helpful independent of learning style, it can be recommended to learners of all types.

It can further be recommended that teachers make an effort to produce vocabulary recordings and make them available to their students and that textbook publishers include audio recordings of their wordlists with the textbooks or on support websites. Materials such as the software Clip2go which combines audio with visual supply of orthographic information, will also prove beneficial. Such resources will likely be appreciated by language learners, who, as we saw above, appear to believe the method to be effective, fun and less work than paired-associates memorisation.

Questions that could not be answered by this study and that would benefit from future research include necessity or otherwise of visual input (orthographic form) as part of an audio-supported approach, what the influence of number and frequency of

exposures to treatment is, how much concentration on the audio recordings is necessary (i.e. would doing housework while listening still yield the same results) and whether results will generalise across different types of language learners in different educational settings. Furthermore, the provisional claim that auditory vocabulary learning is superior due to the nature of verbal processing will need to be examined in more detail.

References

Anon. (1987) *VocabuLearn Korean Level 1*. Carlsbad, CA: Penton Overseas, Inc

Baddeley, A. D. and G. Hitch (1974) Working memory. *The Psychology of Learning and Motivation* 8: 47-89.

Baddeley, A. D. (2003) Working memory and language. *Journal of Communication Disorders* 36(3): 189-208.

Baddeley, A. D. (2007) *Working Memory, Thought, and Action*. Oxford: OUP

Bandler, R. and J. Grinder (1979) *Frogs into Princes: Neuro Linguistic Programming*. Moab, UT: Real People Press

Bird, S. and G. Jacobs (1999) An examination of the keyword method: How effective is it for native speakers of Chinese? *Asian Journal of English Language Teaching* 9: 75-97

Cho, H. (2002) The effect of vocabulary exercise types on L2 vocabulary retention. *SNU Working Papers in English Linguistics and Language* 1: 204-23.

Clip2go Basic Vocabulary. [digital audio player software] (2006) ClipTip, Ohmstrasse 14, 80802 Munich, Germany, cliptip.de

Cohen, J. (1988) *Statistical Power Analysis for the Behavioural Sciences*. Hillsdale, NJ: Lawrence Earlbaum Associates.

Council of Europe (2001) *Common European Framework of Reference for Languages*. Cambridge: Cambridge University Press

de Groot, A. (2006) Effects of stimulus characteristics and background music on foreign language vocabulary learning and forgetting. *Language Learning* 56 3: 463-506.

de la Fuente, M. (2006) Classroom L2 vocabulary acquisition: investigating the role of pedagogical tasks and form-focused instruction. *Language Teaching Research* 10 3: 263-95.

Ellis, R. (1994) *The Study of Second Language Acquisition*. Oxford: OUP

Français Interactif: Liste de Vocabulaire (2007) Podcast [Online] available from http://www.laits.utexas.edu/fi/html/voc/00.html [Accessed 1 October 2007 14:33:00 GMT]

Gillard, P. (senior commissioning ed) (2003) *Cambridge Advanced Learner's Dictionary, 2nd edition*. Cambridge: Cambridge University Press.

Grinder, M. (1991) Righting the Educational Conveyor Belt. Portland, OR: Metamorphous Press.

Hedges, L. and I. Olkin (1985) *Statistical Method for Meta-Analysis*. London: Academic Press

Henriksen, B. (1999) Three dimensions of vocabulary development. *Studies in Second Language Acquisition* 21: 303-17

Hermann, F. (2003) Differential effects of reading and memorization of paired associates on vocabulary acquisition in adult learners of English as a second language. *TESL-EJ* 7 1: 3-15.

Just Vocabulary (2007) Podcast [Online]. available from http://www.wizzard.tv/ shows2/popups/justvocabulary [Accessed 1 October 2007 14:28:00 GMT]

Laufer, B. (1997) What's in a word that makes it hard or easy: some intralexical factors that affect the learning of words. In N. Schmitt and M. McCarthy (eds.). pp. 140-55.

Lee, A. (2001) Korean speakers. In M. Swan and B. Smith (eds.) *Learner English*. Cambridge: Cambridge University Press. pp. 325-42.

Lotto, L. and A. de Groot (1998) Effects of learning method and word type on acquiring vocabulary in an unfamiliar language. *Language Learning* 48: 31-69.

Meara, P. and G. Jones (1990) *Eurocentres Vocabulary Size Tests 10*. Zurich: Eurocentres Learning Service.

Melka, F. (1997) Receptive vs. productive aspects of vocabulary. In Schmitt and McCarthy 1997: 84-102

Nation, I. S. P. (1990) *Teaching and Learning Vocabulary*. New York: Newbury House.

Papagno, C., T. Valentine and A. Baddeley (1991) Phonological short-term memory and foreign-language vocabulary learning. *Journal of Memory and Language* 303: 331-47.

Paribakht, T. and M. Wesche (1993) The relationship between reading comprehension and second language development in comprehension-based ESL program. *TESL Canada Journal* 111: 9-29

Paribakht, T. and M. Wesche (1996) Enhancing vocabulary acquisition through reading: a hierarchy of text-related exercise types. *Canadian Modern Language Review* 52: 250-73.

Pitts, M., H. White and S. Krashen (1989) Acquiring second language vocabulary through reading: a replication of the Clockwork Orange study using second language acquirers. *Reading in a Foreign Language* 5 2: 271-75.

Rott, S., J. Williams and R. Cameron (2002) The effect of multiple-choice L1 glosses and input-output cycles on lexical acquisition and retention. *Language Teaching Research* 63: 183-222.

Sagarra, N. and M. Alba (2006) The key is in the keyword: L2 vocabulary learning methods with beginning learners of Spanish. *The Modern Language Journal* 90 ii: 228-43.

Schmitt, N. (1997) Vocabulary learning strategies. In Schmitt and McCarthy (eds.) pp. 199-227

Schmitt, N. and M. McCarthy (eds.) (1997) *Vocabulary: Description, Acquisition and Pedagogy*. Cambridge: CUP

Service, E. (1992) Phonology, working memory and foreign-language learning. *The Quarterly Journal of Experimental Psychology* 45(1): 21-50.

Shapiro, A. and D. Waters (2005) An investigation of the cognitive processes underlying the keyword method of foreign vocabulary learning *Language Teaching Research*, 9 (2): 129-46.

Ulanoff, S. and S. Pucci (1999) Learning words from books: the effects of read aloud on second language vocabulary acquisition. *Bilingual Research Journal*, 23 (4): 319-32.

Vidal, K. (2003) Academic listening: a source of vocabulary acquisition? *Applied Linguistics* 24 1: 56-89

Vocabulario Español (2007) Podcast [Online] available from http://techteacher.libsyn.com/ [accessed 10 October 2007 14:37 GMT]

Waring, R. and M. Takaki (2003) At what rate do learners learn and retain new vocabulary from reading a graded reader? *Reading in a Foreign Language* 15 2: 130-63.

Webb, S. (2007) Learning word pairs and glossed sentences: the effects of a single context on vocabulary knowledge. *Language Teaching Research* 111: 63-81.

Wesche, M. and S. Paribakht (1996) Assessing second language vocabulary knowledge: Depth versus breadth. *The Canadian Modern Language Review*, 53: 13-40

Wingate, J. (2000) Learning preferences. *English Teaching Professional*, 17: 31-2.

Witkin, H.A., P. Oltman, E. Raskin and S. Karp (1971) *A Manual for The Group Embedded Figures Test*. Palo Alto: Consulting Psychologists Press

Witkin, H.A., C. Moore, D. Goodenough and P. Cox (1977) Field dependent and field independent cognitive styles and their educational implications. *Review of Educational Research* 47:1-64.

14

Appendix A: Treatment handouts (used during all four treatment sessions)
Group 1

SET 1 (group 1) - auditory mode

gravel (모래사 섞인) 자갈

primary 가장 중요한

acquiesce 순순히 받아들이다, 따르다

batch (물건의) 한 묶음

connoisseur (미술이나 음악에 있어) 전문가

jumble 뒤죽박죽 된 것

whiz 매우 빨리 지나가다

consent 승낙, 동의

crease 주름지다, 주름을 만들다

smoulder (불꽃을 내지 않고) 천천히 타다

SET 2 (group 1) - non-auditory mode

bamboozle
/bæm'bu:zl/ ...를 속이다, 끌탕먹이다
ex: She was bamboozled into tell him her credit card number.
ex: Some banks try to bamboozle their customers.

obsolete
/ˌɒb.sə'li:t/ 구식의
ex: Gas lamps became obsolete when electric lights were invented.
ex: Most computers become obsolete a few months after you buy them.

revelry
/'rev.əl.ri/ 흥겹고 시끄럽게 놀기
ex: Sounds of revelry came from next door.
ex: The revelries next door kept me awake all night

antics
/'æn.tɪks/ 우스꽝스러운 동작
ex: The rock-star's stage antics used to include smashing guitars.
ex: The crowds were once again entertained by the tennis player's antics.

dregs
/dregz/ 찌꺼기
ex: In one swift go, she had drunk her coffee down to the dregs.
ex: When you finish drinking a cup of coffee, the dregs are left at the bottom of the cup.

blue-blooded
/blu:'blʌd.ɪd/ 귀족 (명문가) 출신 의
ex: She was not blue-blooded, but she was an officer's daughter.
ex: Jenny's fiancé is actually blue-blooded.

bask
/bɑːsk/ 햇볕을 쬐다
ex: We sat basking in the warm sunshine.
ex: A cat was basking on the windowsill.

skid
/skɪd/ 미끄러지다
ex: Trevor's bus skidded on some ice and hit a tree.
ex: Her foot skidded on the wet floor and she fell hard.

vacuous
/'væk.ju.əs/ 생각이 결여된, 얼빠진
ex: John and his vacuous fiancée will come for dinner tonight.
ex: His father is always ready with vacuous advice.

seclusion
/sɪ'klu:.ʒən/ 격리, 차단
ex: After the tour I was happy to return to the seclusion of my own home.
ex: Today there are 25 monks living in seclusion at Highbury monastery.

Group 2

SET 1 (group 2) - non-auditory mode

gravel
/ˈɡrævəl/
(모래사섞인) 자갈
ex: There is a gravel path leading to the house.
ex: There were few vehicles on the gravel roads.

primary
/ˈpraɪ.mə.ri/
가장 중요한
ex: The Red Cross's primary concern is to preserve and protect human life.
ex: The primary reason for the bad economic situation is government policy.

acquiesce
/ˌæk.wiˈes/
순순히 받아 들이다, 따르다
ex: Reluctantly, he acquiesced to the plans.
ex: In the end, the minister acquiesced because there was no other way.

batch
/bætʃ/
(물건의) 한 묶음
ex: The cook brought in a fresh batch of homemade cakes.
ex: We looked at job applications in two batches.

connoisseur
/ˌkɒn.əˈsɜːr/
(미술이나 음악에 있어) 전문가
ex: Aaron is a wine connoisseur
ex: A connoisseur of cigars will know the difference.

jumble
/ˈdʒʌm.bl̩/
뒤죽박죽 된 것
ex: He tried to find the letter among the jumble of papers on his desk.
ex: The essay was a meaningless jumble of ideas.

whiz
/wɪz/
매우 빨리 지나가다
ex: A police car whizzed by on its way to the accident.
ex: The time just whizzes past when you are enjoying yourself.

consent
/kənˈsent/
승낙, 동의
ex: They can't publish your name without your consent.
ex: You can only come on the trip if your parents give their consent.

crease
/kriːs/
주름지다, 주름을 만들다
ex: The seatbelt has creased by shirt.
ex: It's a nice dress, but it creases easily.

smoulder
/ˈsməʊl.dər/
(불꽃을 내지 않고) 천천히 타다
ex: The bonfire was still smouldering the next day.
ex: The cigarette was smouldering when he put it in the bin. That's how the fire started.

SET 2 (group 2) - auditory mode

bamboozle
...을 속이다, 골탕먹이다

obsolete
구식의

revelry
흥겹고 시끄럽게 놀기

antics
우스꽝스러운 동작

dregs
찌꺼기

blue-blooded
귀족 (명문가) 출신 의

bask
햇볕을 쬐다

skid
미끄러지다

vacuous
생각이 결여된, 얼빠진

seclusion
격리, 차단

16

Appendix B: form used for 48-hour post-test (t1), phase 1 (identical forms were used for t2 and t3)

immediate post (48hrs)

student number: _____

time allowed: 15 minutes

지시사항
- 사전이나 다른 도움없이 하시기 바랍니다.
- A)에서는 설명을 읽으신후 빈칸에 그뜻을 영어단어로 적어주십시오.
- B)에서는 그 단어를 사용하여 짧은 한 문장을 만들어 주십시오. 단 예제로 주어진 목록에 있는 문장을 똑같이 그대로 사용하지 말아 주십시오.

예:
A) 크래식 classical
B) I love classical music.

1 A) 순순히 받아 들이다, 따르다 _____
B) _____

2 A) 미끄러지다 _____
B) _____

3 A) 주름지다, 주름을 만들다 _____
B) _____

4 A) 생각이 결여된, 얼빠진 _____
B) _____

5 A) 승락, 동의 _____
B) _____

6 A) 귀족 [명문가] 출신 의 _____
B) _____

7 A) (불꽃을 내지 않고) 천천히 타다 _____
B) _____

8 A) 째개기 _____
B) _____

9 A) 티죽박죽 된 것 _____
B) _____

10 A) 거리, 처단 _____
B) _____

11 A) (모래사 섞인) 자갈 _____
B) _____

12 A) 햇볕을 쬐다 _____
B) _____

13 A) (물건의) 한 묶음 _____
B) _____

14 A) 즐겁고 시끄럽게 놀기 _____
B) _____

15 A) 매우 빨리 지나가다 _____
B) _____

16 A) 우스꽝스러운 동작 _____
B) _____

17 A) 가장 중요한 _____
B) _____

18 A) 구석의 _____
B) _____

19 A) (미술이나 음악에 있어) 전문가 _____
B) _____

20 A) ...을 속이다, 골탕먹이다 _____
B) _____

Table 1: Combinatory matrix of groups, modes and sets

	auditory mode	non-auditory mode
group 1	set 1	set 2
group 2	set 2	set 1

Table 2: Information supplied and channels used in auditory and non-auditory modes

	auditory mode	non-auditory mode
orthographic form of target	visual only	visual only
L1 translation/explanation	visual only	visual only
phonological form	auditory only	visual only (IPA-transcription)
two example sentences	auditory only	visual only

```
┌─────────────────┐      ┌─────────────────┐
│  visual input   │      │ auditory input  │
└─────────────────┘      └─────────────────┘
         │                        │
         ▼                        ▼
┌─────────────────┐      ┌──────────────────────────────┐
│ visual   short- │      │  ┌─────────────────┐    │ ph  │
│ term store      │      │  │ phonological    │    │ on  │
└─────────────────┘      │  │ short-term      │    │ ol  │
         │               │  │ store           │    │ og  │
         ▼               │  └─────────────────┘    │ ic  │
┌─────────────────┐      │      │        ▲         │ al  │
│ orthographic to │      │      ▼        │         │ lo  │
│ phonological    │─────▶│  ┌─────────────────┐    │ op  │
│ recording       │      │  │ articulatory    │    │     │
└─────────────────┘      │  │ rehearsal       │    │     │
                         │  │ mechanism       │    │     │
                         │  └─────────────────┘    │     │
                         └──────────────────────────────┘
```
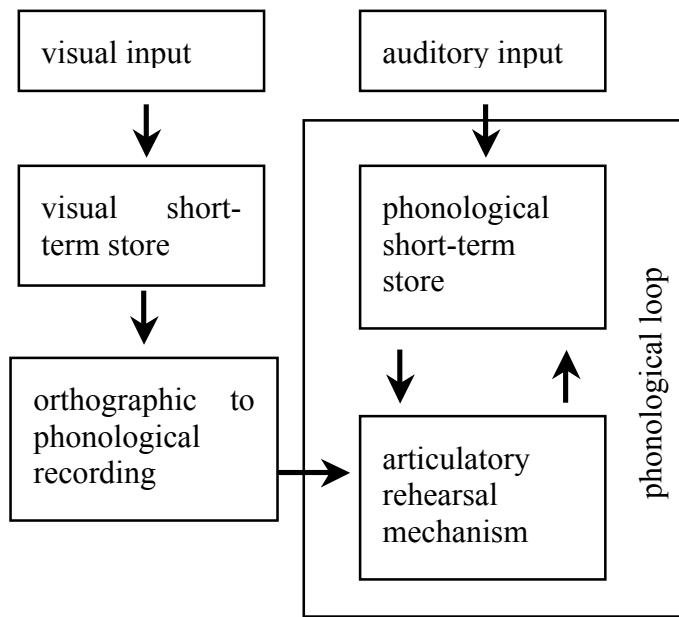
Figure 1: Processing of visual and auditory input in working memory (based on Baddeley 2003: 193)

pre-test (selection of test items)

↓

4 treatment sessions within 8 days each treatment session lasted 14 minutes 7 minutes for set 1 7 minutes for set 2

↓

48-hr post-test (t1): all criteria tested

↓

10-day post-test (t2): written test only (pronunciation not tested)

↓

2-month post-test (t3): all criteria tested
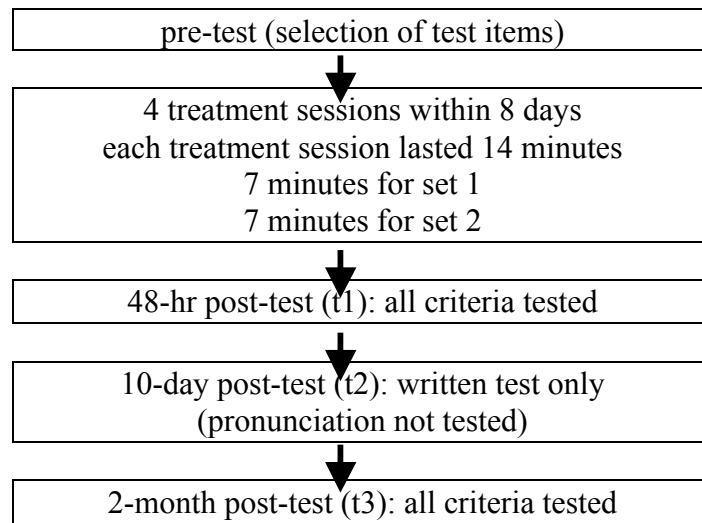
Figure 2: Overview of procedure

Table 3: Raw and mean scores for test items by set for t1 (n=62)

set 1			set 2		
word	total	mean	word	total	mean
acquiesce	103	1.66	revelry	129	2.08
connoisseur	133	2.15	antics	154	2.48
whiz	168	2.71	obsolete	157	2.53
jumble	177	2.85	vacuous	183	2.95
consent	181	2.92	seclusion	185	2.98
batch	187	3.02	dregs	195	3.15
smoulder	196	3.16	bamboozle	200	3.23
gravel	211	3.40	skid	226	3.65
crease	229	3.69	bask	230	3.71
primary	246	3.97	blue-blooded	280	4.52
means set 1	183.1	2.95	means set 2	193.9	3.13
p = 0.582					

Note: Means were calculated by dividing the total score by n. The significance of differences between sets was calculated using a two-tailed t-test.
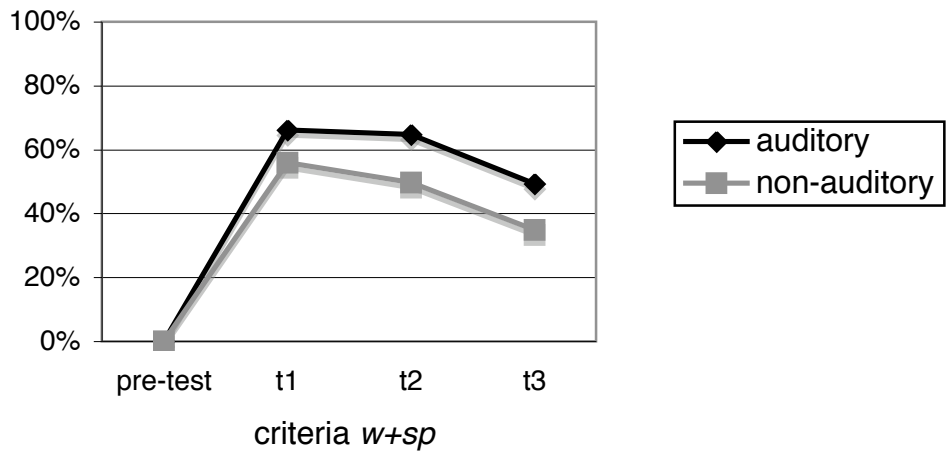
Figure 3: Scores for criteria w+sp in percentages of maximum scores across pre-test, t1, t2 and t3.

Table 4: Mean scores (standard deviations), effect sizes and significance

		Mean auditory score (sd)	Mean non-aud. score (sd)	MD (mean difference)	unbiased d	p
t1						
w	62	7.1 (2.2)	5.9 (2.8)	1.2	0.46	0.012
w+sp	62	13.2 (4.5)	11.2 (5.3)	2.0	0.41	0.023
w+gr	62	11.0 (3.4)	8.9 (4.5)	2.1	0.51	0.005
w+sm	62	11.3 (3.7)	9.2 (4.6)	2.1	0.49	0.007
w+sp+gr+sm	62	21.4 (7.3)	17.5 (8.9)	3.9	0.47	0.009
st	45	8.5 (1.2)	7.5 (1.2)	1.0	0.89	0.000
phon	45	8.2 (1.4)	6.4 (1.9)	1.8	1.07	0.000
st+phon	45	16.7 (2.1)	13.8 (2.6)	2.9	1.21	0.000
overall	45	37.4 (6.4)	32.6 (10.2)	4.8	0.57	0.008
t2						
w	62	6.9 (2.3)	5.4 (2.8)	1.5	0.61	0.001
w+sp	62	13.0 (4.5)	10.0 (5.2)	3.0	0.61	0.001
w+gr	62	11.3 (3.9)	8.4 (4.4)	2.9	0.69	0.000
w+sm	62	11.7 (4.0)	8.7 (4.5)	3.0	0.70	0.000
w+sp+gr+sm	62	22.1 (7.9)	16.3 (8.6)	5.8	0.70	0.000
t3						
w	62	5.3 (2.1)	3.7 (2.3)	1.6	0.73	0.000
w+sp	62	9.9 (4.0)	7.0 (4.4)	2.9	0.69	0.000
w+gr	62	8.5 (3.8)	6.0 (3.7)	2.5	0.66	0.000
w+sm	62	8.6 (3.8)	6.0 (3.7)	2.6	0.70	0.000
w+sp+gr+sm	62	16.4 (7.4)	11.6 (7.1)	4.8	0.66	0.000
st	45	8.3 (1.3)	7.2 (1.4)	1.1	0.77	0.000
phon	45	7.8 (1.4)	6.8 (1.6)	1.0	0.66	0.002
st+phon	45	16.1 (2.2)	14.0 (2.1)	2.1	0.95	0.000
overall	45	31.6 (7.8)	25.3 (8.2)	6.3	0.78	0.000
difference between t1 and t3 (forgetting)						
w	62	1.7 (2.0)	2.2 (2.2)	-0.5	-0.22	0.219
w+sp	62	3.4 (3.8)	4.2 (4.4)	-0.8	-0.21	0.238
w+gr	62	2.4 (3.4)	2.9 (3.6)	-0.5	-0.12	0.486
w+sm	62	2.7 (3.5)	3.2 (3.4)	-0.5	-0.16	0.367
w+sp+gr+sm	62	5.0 (6.7)	5.9 (6.8)	-0.9	-0.14	0.443
st	45	0.2 (1.5)	0.2 (1.6)	0.0	0.01	0.947
phon	45	0.4 (1.8)	-0.4 (2.3)	0.8	0.39	0.067
st+phon	45	0.6 (2.7)	-0.2 (2.9)	0.8	0.29	0.170
overall	45	5.8 (7.3)	7.3 (7.3)	-1.5	-0.20	0.352

Note: Unbiased d is Cohen's d with Hedges and Olin's (1985) adjustment. Numbers involving the *st* and *phon* criteria are lower because only a random subset of participants (the same subset for all tests) were tested on these criteria. Scores for groups 1 and 2 were summed as they served as their own controls. T1 is the immediate post-test, t2 the 10-day post-test and t3 the 2-month post-test. 'Overall' is *w+sp+gr+sm+st+phon*. *p*-values for MD were calculated using two-tailed t-tests (alpha = .05)

Table 5: Analysis by participants

	n	participants with higher auditory scores	participants with higher or equal auditory scores	participants with lower auditory scores
t1				
w	62	56.5%	72.6%	27.4%
w+sp	62	58.1%	64.5%	35.5%
w+gr	62	59.7%	74.2%	25.8%
w+sm	62	62.9%	74.2%	25.8%
w+sp+gr+sm	62	59.7%	69.4%	30.6%
st	45	66.7%	84.4%	15.6%
phon	45	68.9%	84.4%	15.6%
st+phon	45	80.0%	80.0%	20.0%
overall	45	73.3%	77.8%	22.2%
t2				
w	62	62.9%	79.0%	21.0%
w+sp	62	67.7%	72.6%	27.4%
w+gr	62	66.1%	72.6%	27.4%
w+sm	62	69.4%	77.4%	22.6%
w+sp+gr+sm	62	66.1%	71.0%	29.0%
t3				
w	62	69.4%	85.5%	14.5%
w+sp	62	74.2%	82.3%	17.7%
w+gr	62	72.6%	83.9%	16.1%
w+sm	62	75.8%	83.9%	16.1%
w+sp+gr+sm	62	77.4%	82.3%	17.7%
st	45	57.8%	75.6%	24.4%
phon	45	60.0%	73.3%	26.7%
st+phon	45	66.7%	86.7%	13.3%
overall	45	84.4%	84.4%	15.6%

Note: *overall* is *w+sp+gr+sm+st+phon*

Table 6: Correctly spelt words as percentage of words recalled in t1 (standard deviation in brackets), n = 62

auditory mode	85.0% (19.7)
non-auditory mode	89.6% (13.9)

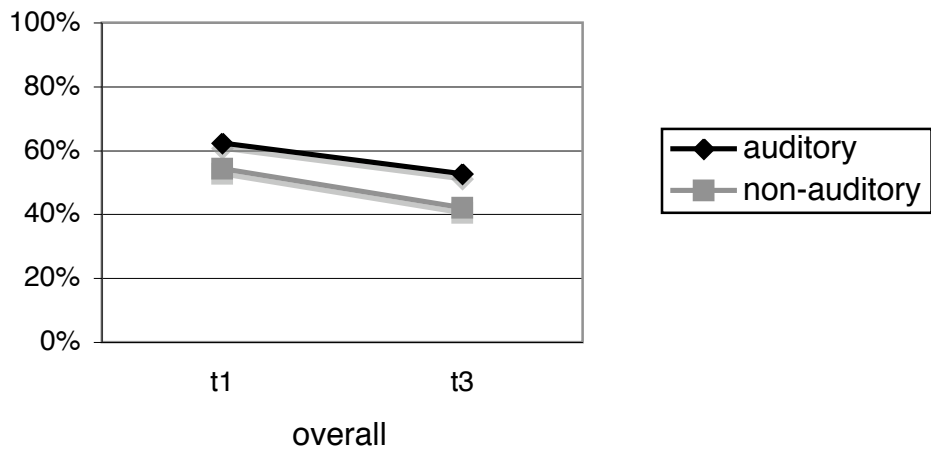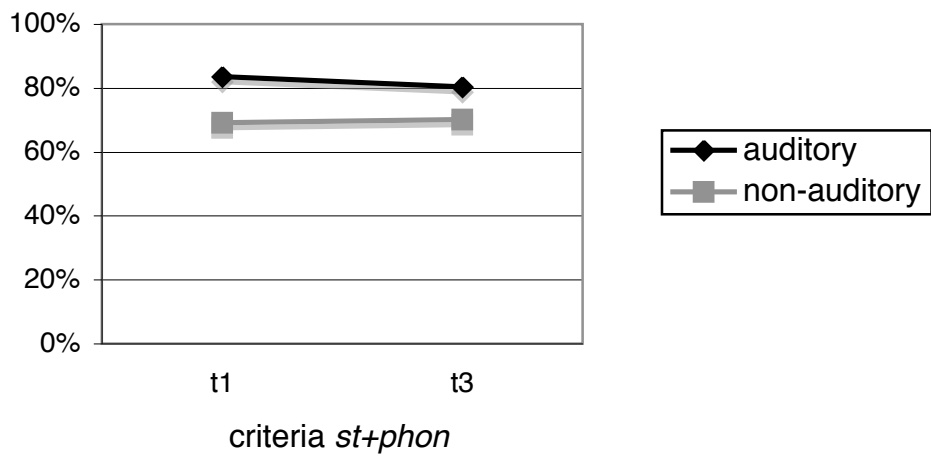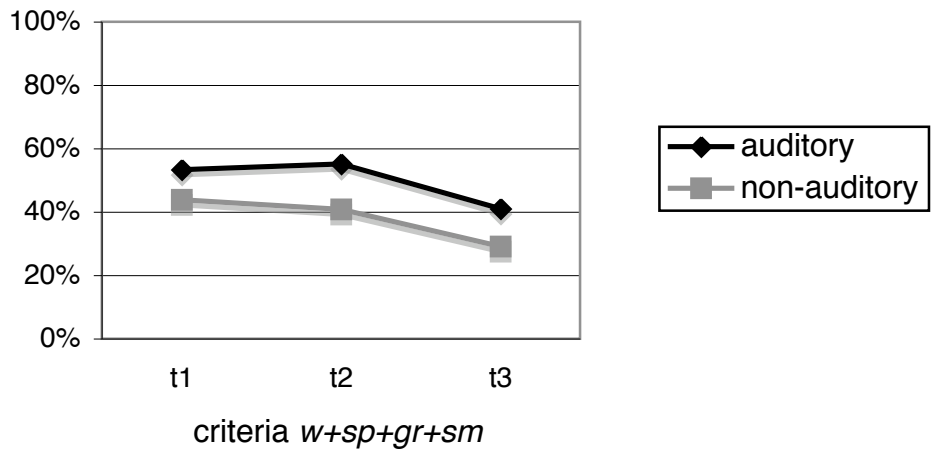*p* = 0.135 (two-tailed t-test for the difference between modes)

Figure 4: Scores for three different configurations of criteria in percentages of maximum scores for each criteria configuration

Table 7: Attitudes toward auditory and non-auditory modes (n = 85)

Questionnaire:

You have now studied words 1) silently and 2) with audio support.

A) Which way was more work?

B) Which way was more effective?

C) Which one was more fun?

D) If you had a choice, which way would you prefer to use to study words in the future?

Answers:	non-auditory	auditory	both/neither
A)	61.2% (52)	35.3% (30)	3.5% (3)
B)	27.1% (23)	72.9% (62)	0.0% (0)
C)	10.6% (9)	88.2% (75)	1.2% (1)
D)	35.3% (30)	54.1% (46)	10.6% (9)

Note: Number of respondents is given in brackets, the questionnaire was administered immediately after the last treatment session.

Table 8: Correlations between learning preference scores and t1 scores

post-tests	n	Learning preference scores auditory		visual		kinaesthetic	
t1		r	*p*	r	*p*	r	*p*
w	62	0.246	>.05	-0.182	>.05	-0.085	>.05
w+sp	62	**0.261**	**0.040**	-0.200	>.05	-0.082	>.05
w+sp+gr+sm	62	0.211	>.05	-0.200	>.05	-0.020	>.05
st+phon	45	0.129	>.05	0.104	>.05	-0.283	>.05
overall	45	**0.332**	**0.026**	-0.205	>.05	-0.164	>.05

Note: Correlations are between learning preference scores and the difference between auditory and non-auditory scores of t1, significant correlations in bold print.
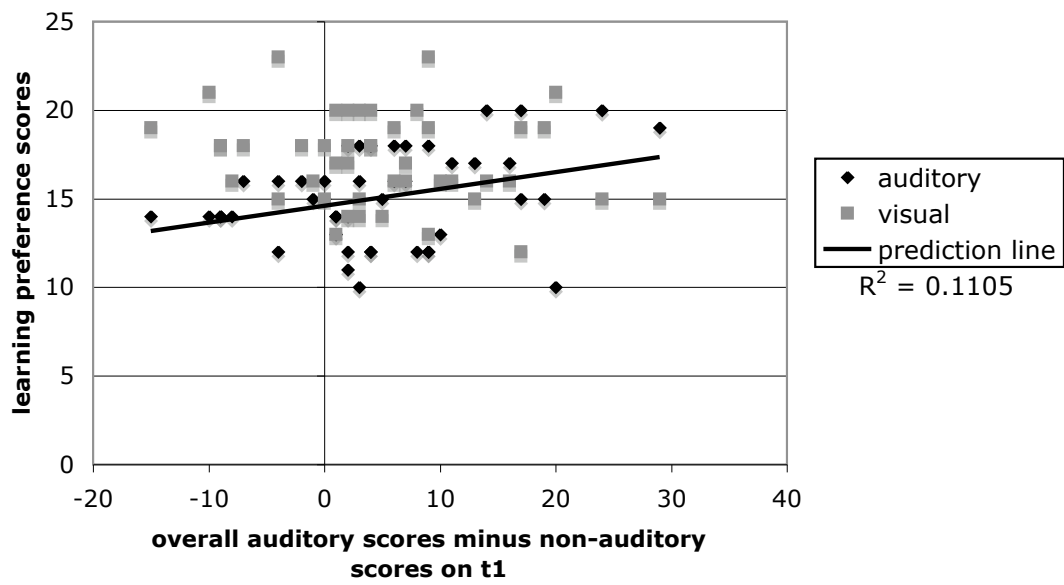
Figure 5: Scatter diagram of correlation between learning preference scores and the difference between auditory and non-auditory scores on t1 (w+sp+gr+sm+st+phon), n = 45; prediction line and R square value for auditory learning preference scores.