

**HOW EFFECTIVE ARE WORDLISTS FOR LISTENING TO
AUTHENTIC FOREIGN LANGUAGE NEWS BROADCASTS?**

Peter Connell

ABSTRACT

Due to the difficulty that Japanese learners may face in understanding the gist of authentic British and American TV broadcasts, their instructors often believe that transcripts from the listening content should be available beforehand. However, there is reason to believe that such assistance does not challenge learners sufficiently to exercise their listening ability for ultimately understanding the gist of a passage. On the other hand, offering no support at all during especially authentic listening tasks could drain learner confidence and motivation. With these assumptions in mind, an experiment was conducted to see if a wordlist given to students would be a more effective alternative that aids learner comprehension without reducing the need to utilize aural skills. The results indicate that this pre-listening activity is inadequate for learners who are required to actively show how well they understand the main idea from authentic listening input. The author suggests that they are often not capable of digesting and making good use of the vocabulary they are given without more extensive practice and constructive exposure.

Key Words: wordlist, listening, gist, authentic

INTRODUCTION

Before language learners are given challenging listening material such as authentic TV news broadcasts, the instructor often has to decide whether or not to offer pre-listening support in the form of transcripts from the video segments. Certainly, there are language instructors who assume that such supportive features are beneficial and even necessary for their students, many of whom may be considered low-level. The

Peter Connell

question arises, however, as to whether transcripts have become so pivotal that they reduce the learners' motivation to become reliant on their own aptitudes for comprehending listening input. Certainly, low-level language learners may need some lexical support, just as new swimmers wear life jackets as they gradually enter the deep end. Having said that, would the offer of a simple list of key words, without the accompanying definitions of a glossary, be a means of striking the balance needed to ensure that learners are provided with significant assistance without being spoon-fed to the point of relying more on their reading abilities than those associated with listening?

To answer this question, an experiment was conducted with two groups of randomly-chosen Japanese university students of relatively equal English language ability to see whether one group with the assistance of a wordlist can comprehend British and American TV news broadcasts better than the one that listens to the broadcasts unaided by this prop. If the two groups show no noteworthy differences in their results, then a conclusion can be drawn that wordlists on their own are not necessarily effectual. However, if the wordlist group does significantly better, then an instructor may feel that the pre-listening procedure of handing out lists of key words plays the role of a sensible balance between the use of transcripts and the provision of no lexical support at all.

This article will summarize some of the reasons behind the poor English language oral skills of Japanese learners, particularly top-down listening, and suggest an effective way to collectively assess their ability to master the gist of short but authentic American/British TV news clips. After this initial foundation is established, an experiment will be carried out that will try to ascertain how much better these learners would do in deciphering authentic listening content if tasks were preceded by aforementioned wordlists.

LITERATURE REVIEW

When providing students with challenging listening material, many instructors/researchers with relevant experience in Japan have utilized the transcripts in a variety of ways for their classes. Obermeir (2000) advocated the 'listening-while-reading-the-transcripts' approach for guiding his students who were studying for the listening segment of the TOEFL. The inevitable goal was to eventually get the students to feel

confident enough to be able to listen without reading. Alberding (2004) carried out a similar approach but for the purpose of having students locate stress patterns in sentences. Decker (2004) provided a summary of self-study listening materials that allow students to do various activities with access to answer keys and scripts. There is little doubt that such methods of using transcripts can aid lexical development, but what is not so clear is why learners would feel compelled to actually use their aural powers when they know that they can avoid doing so in favour of a task for which they have formally had plenty more experience — reading.

At this point, the poor listening skills in relation to reading skills of Japanese learners of English require elaboration in order to provide perspective as to why many language teachers may feel they have no choice but to rely on the use of transcripts to get students through communicative listening assignments. First of all, in terms of formal education in Japan, an imbalance presently exists in which written materials far outweigh listening materials in language study (Fujimoto, 1999; Matsuya, 2003; Shimo, 2002). Even when these listening features are employed, the tendency to overemphasise phonology, grammatical structure and memorisation has not been conducive for students wanting to understand English that is naturally spoken (Suenobu, 1989). These students become “like an oscilloscope or a sound spectrograph” (p. 7). This strategy of listening for sounds as opposed to meaningful communication is an indication of how far students have to go in mastering the language. Therefore, their strategy for listening, reinforced by the method of teaching, produces unsatisfactory results.

Suenobu (1995) wonders if there are a large number of Japanese learners of English “who believe that ‘listening’ is not needed for mastering the English language” (p. 113). This worry may be exaggerated, but it often appears to be the case that the English language curriculum in Japanese grade schools have tended to pay particular attention to correct grammar that students use with certain words that must be produced rather than whether learners actually understood the content of what was spoken (Tajima, 1978). Such intolerance and opposition to carrying out flexible approaches to testing learners’ listening comprehension is reminiscent of circus performers forcing animals to jump through hoops as well as doing other unnatural acts.

Undoubtedly, Japanese learners with a background associated with commonly-taught teacher-centred memorisation tasks would have enormous difficulty in adjusting to the initial phases of understanding the

Peter Connell

main ideas of natural conversations. Dunkel (1986), speaking generally about language learners being initially exposed to non-contrived listening material, describes the experience in this way:

With nonscripted, ungraded “authentic” language tapes, certain difficulties surface which may, in fact, cause learner frustration and demoralization. It is particularly difficult for the beginning-level student to disentangle the thread of discourse, to identify different voices, and to cope with frequent overlaps in segments of authentic language presented via audiotape. (p. 101)

Therefore, the bulk of Japanese language learners need to make a complete reversal in many of their previously-ingrained learning habits as they adopt a more holistic approach to listening exercises. By doing this, they derive meaning based on prior knowledge, otherwise known as top-down listening (Richards, 1990), and would replace the emphasis on ‘bottom-up’ micro-skills of processing listening input in which the learner decodes sounds of a language into words, clauses and sentences (Richards, 2003).

Essentially, an evaluation system of a learner’s ability to comprehend an oral passage needs to be able to distinguish whether the learner actually misunderstood listening passages or whether he/she just poorly communicated the answers. In addition, the examiner has to use discretion to decide the severity of errors that obstruct reader comprehension of learner responses.

Even when a suitable means of evaluation of learners is devised in order to ascertain their ability to understand the gist of a video segment, a certain degree of reliability is also required. Some methods of assessment may demonstrate findings that are objectively acquired, but provide little insight into how well learners understood the answers to questions. Questions falling into this category often require answers in the form of ‘Yes/No’, ‘filling in the gap with a word’ or ‘multiple choice’, in which correct answers can often be obtained through the process of elimination. Occurrences such as these would be to the detriment of construct validity (Wu, 1998). There is no evidence that learner responses to such tests show comprehension (Brown & Yule, 1983), and therefore the reliability of this approach is severely impeded.

The more opportunity that learners have to express what exactly they understand, the easier it is for the examiner to judge their degree of success.

Therefore, to increase the diagnostic potential of a listening-for-gist exercise, it is preferable to have learners write responses in full-sentence answers, bearing in mind that the information not only identifies correctness or error, but also may provide clues as to whether these learners had simply poorly expressed what they may have correctly heard. As a result, the value of the research findings will depend not only on whether responses are correct or not, but also on whether both the question and answer were actually understood by the learner.

Inevitably, the requirement of Japanese learners to write full sentences for their responses is a challenging endeavour when one considers their tendency to lack confidence in their communicative foreign language skills. Freiermuth (1997) feels such students should be commended for attempting to push the boundaries of their language abilities. Certainly, building learner confidence is a high priority in administering language activities, but at what level of strictness do errors need to be pointed out?

According to South (1999), students who are learning general English for the purpose of improving their overall proficiency should perhaps only have their errors corrected when they impede understanding. Freiermuth (1997) summarizes his formula for locating serious errors:

It may be useful to view errors in a hierarchy, ranked according to their seriousness, with errors that significantly impair communication at the top of the list, followed by errors that occur frequently, errors that reflect misunderstanding or incomplete acquisition of the current classroom focus, and errors that have a highly stigmatizing effect on listeners. (para. 10)

Holistic evaluations carried out by the instructor alone would appear to be sufficient for providing him or her with the necessary input on where learner problems lie in understanding or expressing the gist of a listening passage. The grading of the learner responses is usually up to the sole discretion of the instructor, who can maintain a consistent monitoring of learner achievement based on set criteria for satisfactory results. The value of this approach is that it focuses assessment mainly on learners' ability to communicate (Jacobs, Zinkgraf, Wormuth, Faye Hartfiel, & Hughey, 1981; Weir, 1990).

Admittedly, the determination of whether a learner has understood the gist of a listening passage is a subjective exercise. A degree of

judgement is called for on the part of the scorer, which will likely deny perfect consistency (Hughes, 1989). Therefore, for the sake of internal reliability, peer instructors, who work with similar profiled groups of learners, can be engaged (Nunan, 1992). These other researchers can take part in evaluating a selection of the same learner responses on the basis of the standardized grading principles in order to verify that learners are assessed consistently and fairly (Hadley & Mort, 1999; Weir, 1990). Once Inter-rater (also known as 'inter-marker') reliability is established, the evaluation of any number of learners of similar profile can be compared.

When the means of providing listening content and assessment can be validated, variables can be tested in search of ways to help learners cope with the challenge of oral English language input. One variable would be to provide learners with a list of key lexical items that are found in a listening passage that they will attempt to comprehend. At the discretion of the instructor, this list of words will be deemed as being potentially unfamiliar to the learners and relevant to the listening context at the same time. An experiment with this variable is meant to gain insight into how well learners can utilise key lexical terms that they are given as a form of listening support.

It is questionable whether providing learners with words on their own, without much guidance or practice on how they are used in communication, can be useful in deciphering English as it is naturally expressed. This view is based on the problems Japanese learners have in decoding authentic oral input that have been previously mentioned as well as the lack of training in using key vocabulary. Furthermore, assuming learners have extensive familiarity with the formal use of these words, gained through years of school reading and writing activities, there is the obstacle of comprehending how they are communicated orally. Guest (1998) views Japanese students as often learning spoken English which is modelled after written forms in the formal education system. He surmises that "learners invariably fail to absorb the nuances of spoken forms with the result that learners often converse as if they were walking textbooks" (para. 4). To show the contrast between written and spoken English, he uses the analogy that spoken forms of language with their looser structures, often regarded as ungrammatical, are viewed as 'poor cousins' of written forms. McCarthy (1991) reinforces this view as he states that speech abounds in verbless clauses, ellipsis such as omitted pronoun subjects and other structures "that would be frowned

upon in “good” writing (p. 143).” However, this distinction in spoken and written forms, while evident in conversations among native speakers of English, may not be so apparent among non-native speakers who are learning English. Therefore, the oral use of words may be especially problematic for learners, Japanese or otherwise, who mainly have been exposed to written styles throughout their formative years in English language education.

Nevertheless, many educators continue to believe a policy of offering such lists with the belief that learners will have the know-how to use them. For example, Stapleton (1998) advocates these lists on their own as a helpful aid pertaining to particular faculties and Griffiee (1997) also regards exercises that involve key words and their definitions as having a beneficial effect of learner success. The problem may arise, however, from the lack of feedback and reinforcement that may be necessary to ensure that this memorisation of words will lead to learners knowing how to utilise them.

Other educators like Daulton (1998) focus on the fact that the Japanese language incorporates a tremendous amount of loanwords from English and, therefore, the assumption is made that Japanese learners already have plenty of lexical resources which they just need to be made aware of. He adds that Japanese students “should learn to have more confidence in their intuitions about new English vocabulary” (Implications for Vocabulary Instruction in Japan, para. 4). While it is true that many of these loanwords are interchangeable with English, the Japanese pronunciation often renders them as sounding like completely different words. For listening exercises, therefore, the advantage of loanwords is severely minimized.

Having outlined the potential problems learners may have in listening to authentic oral speech and using key vocabulary they are given in writing, a listening experiment has been devised which sets out to reinforce the following hypothesis:

Providing a list of key words to a group of Japanese university language learners before they listen to authentic English-language TV news broadcasts will not actually help them obtain better results in answering questions based on the gist of the content.

In the next section, the experiment based on a holistic evaluation of learner output is outlined.

Peter Connell

MATERIALS AND PROCEDURES

Participants

A group of 40 second-year, non-English majors at a Japanese university were randomly chosen as a control group that would compare its success in mastering listening-for-gist tasks with another similarly-profiled group of the same number of participants who did the same tasks a year later. The main difference between the groups, however, is the addition of a list of key words from the video segments that had been given to the latter group as a pre-listening tool.

Listening Content

The listening material was made up of video segments that were split into three parts of approximately 30 seconds each and viewed from a custom-built website created by the instructor and author of this report. The shortness of the listening passages was deemed appropriate in order for answers to questions based on the main idea of the story to be digestible for participants who, on the whole, have limited experience in the use of authentic English. The choice of the news topic was based on certain criteria:

- The topic is of a general or universal nature.
- The topic is aided by contextual data.
- The topic is non-Japanese in origin.
- The topic is relevant to learners for motivational purposes.
- The topic needs to have certain appeal for learners.
- The topic content must accurately reflect comprehension questions.
- The topic content and its use must adhere to legal guidelines.¹

Two video themes were chosen for the wordlist experiment. The first video entitled 'The College Experience' was used to ensure that two groups, the 'Control' group ($n = 40$) and the 'Experimental' group ($n = 40$) were comparable with each other in terms of their English listening-for-gist abilities. Both groups were evaluated without a wordlist involved in order to note whether significant, pre-experimental differences existed among the two groups in their overall success in completing the task. Once it was established that discrepancies were insignificant, then another video entitled 'Do we need the Kyoto Protocol' was chosen to determine

whether the Experimental Group, armed with a wordlist this time, would show improvement in its results in contrast with the 'Control' group.

Website

On the basis of the aforementioned guidelines, the participants were referred to a structured website in which video segments were to be:

- 1) easily accessed
- 2) reasonably clear in terms of sound and vision
- 3) given contextual support through pre-listening exercises
- 4) accompanied by listening comprehension questions
- 5) connected to a means of providing responses and receiving feedback

The following picture (Figure 1) shows a sample of a website page that participants for this case study have used to fulfill 'listening-for-gist' tasks.

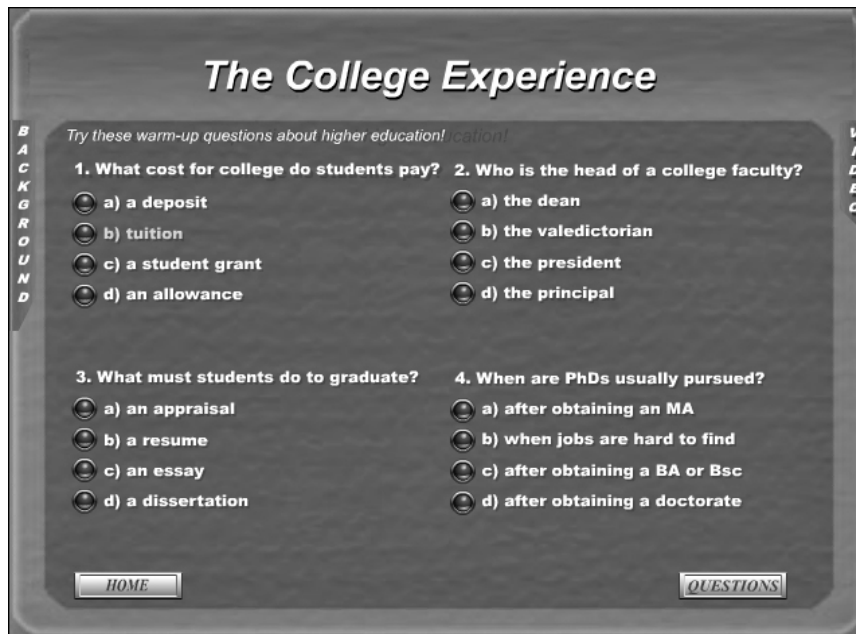


Figure 1. Sample of the Website Home Page and Features

Peter Connell

Wordlist

Before viewing the second video theme (i.e. 'Kyoto'), a sheet of paper with key words from the video was given to the Experimental Group to study and prepare as a pre-listening feature. The vocabulary was chosen on the basis of its high relevance to the listening comprehension questions as well as its potential for being unfamiliar to the participants (see Appendix A for the wordlist).

Procedures

Procedures for using the website

Procedures for the first video theme, 'College', did not include a wordlist. Participants were advised to do the pre-listening warm-up questions which come in a multiple-choice format and provide instant feedback to answers. This feature provided some useful information related to the theme as well as some insight into the video segment that would be accessed soon afterwards. General background information for the upcoming video segment was also provided. After the pre-listening warm-up, the video can be accessed. The panel slides to the left and contains a screen (320x240 pixels) along with a media controller for learners to operate the video (see Figure 2). They encounter an embedded written question before each of the three short (approx. 30 seconds) TV news broadcasts. These questions are specifically designed to be based on the gist of the video segments and have a layout that allows learners to focus on what they need to learn from each video segment instead of having all the questions asked after the viewing. This procedure requires participants to be able to critically assess the information needed to answer the question and essentially dispense with the rest. It was not necessary for the participants to respond with the actual words from the video, but rather to attempt to simply provide a coherent and clear answer based on the video in full-sentence form. The questions for two topics that are used in the experiment are located in Appendix B.

One may ask why it is necessary for participants to answer specific questions at all since the main task is to get them to demonstrate their understanding of the gist. Why not just ask them to summarise the main idea of each video segment? The problem that can be anticipated with that approach is that the participants may have a propensity to provide

vague responses that don't deal with the substance of the listening content. They may still be in need of a certain amount of guidance in learning how specific their answers should be.

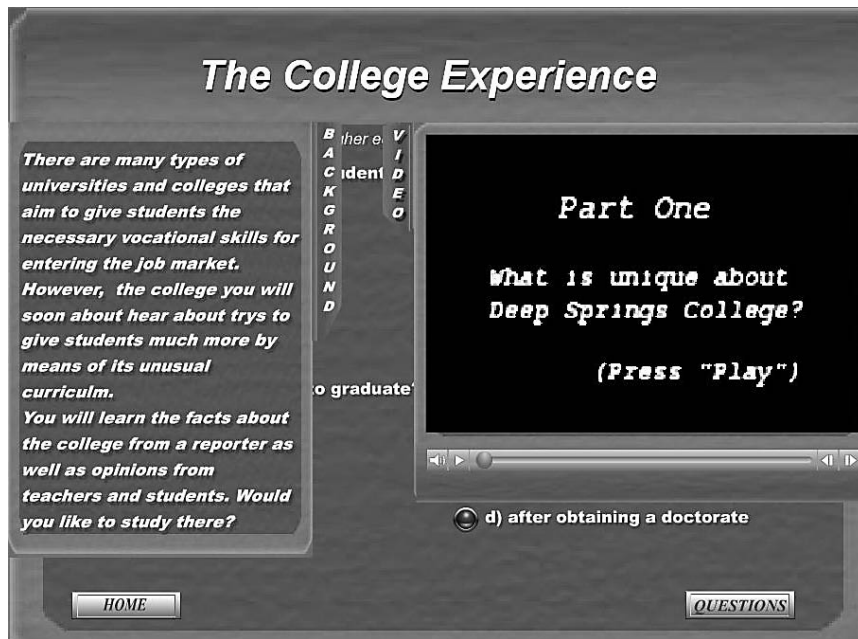


Figure 2. Layout for Pre-Listening Information and the Video

After reading and understanding the questions, the video segments can be viewed repeatedly. However, this freedom is by no means an advantage. It may help to listen to something twice or three times, but learners, themselves, can attest to the fact that if a passage is still hard to understand after that, then no amount of extra viewings is going to be beneficial.

By means of the 'Questions' tab, participant answers can be submitted on the Website Answer Form as seen in Figure 3 below. The top row of boxes elicits participants' personal information and the boxes for question Parts 1, 2 and 3 are positioned to provide response data in the form of one or two full sentences. Participants are then asked to give their idea or opinion in response to a question. This last task originally had nothing to do with the listening-for-gist analysis to come later and was simply a means for learners to exercise a freer (and probably more

Peter Connell

enjoyable) form of writing.

The image shows a web form titled "Website Answer Form" with a dark background. At the top, there are three input fields labeled "Name:", "E-mail Address:", and "Webpage Topic:". Below these are three sections for questions: "Part 1 (What is unique about Deep Springs College?)", "Part 2 (What do most of the students do after graduation?)", and "Part 3 (What do they gain from their experience?)", each followed by a text input area. A question asks, "Any problems understanding the video? (If not, then please type 'No'.)". Below that, a comment section asks, "Comment: Please give an answer to the question below (50+ words). What would you like to learn during your university education and why?" with a large text area. At the bottom are "Submit" and "Reset" buttons.

Figure 3. Sample of a Form for Students to Fill in and Submit

The responses are submitted to the instructor's mail box for analysis and correction, and are subsequently returned by e-mail to the participants.

Procedures for assessment

The instructor as classifier has the basic task of assessing whether sufficient information exists to label a response as correct. If there are serious problems, he/she will indicate what needs to be changed and what lower classification applies. The five category explanations with actual examples of participant answers are distinguished below. The examples are in response to the first question from the theme, 'College' (see Table B1 of Appendix B).

Correct responses. These responses demonstrate knowledge of the main premise of the question, without major lexical/grammatical obstructions or the addition of details that are misleading in terms of contradicting the

video content. For example, in response to the question, ‘What is unique about Deep Springs College?’, the video segment can be transcribed as follows: “Deep Springs College combines intense academics with hard labour.” Participants can use their own words to conform to this answer under these conditions:

- a) *Articulated basic idea from video despite lexical/grammatical problems*
“This college combines classes with hard labor as educational experience.”
- b) *Basically correct with non-contradictory irrelevance*
“The college is located in the desert and the students experience hard labor in order to study for better life.”

Partially-correct responses. Some relevant knowledge of the answer is displayed but clarity is marred either lexically, grammatically or in terms of content. There may also be some false details in the response that are misleading.

- a) *Some relevance to the question but important aspect missing*
“In the college, students can have the experience about agricultural work, really.”
- b) *Some relevant points but details are misleading or off-topic*
“The object they learn is not a text but nature. And they deep discuss about nature, seem to have enthusiastic.”

Incorrect responses. The answer given does not reflect the video content in any way.

- a) *Not relevant due to excessive vagueness*
“Students study by experience.”
- b) *Contradictory in terms of the main idea*
“The college began to educate about dairy.”

Incomprehensible responses. Incoherent script is written that renders the response as impossible to decipher. This may be due to the participant’s limited grammatical awareness or simply a poor attempt to copy every word from a listening passage.

- a) *Poorly comprehended key words*
“It is an educational experience combining in tenth academics.”

Peter Connell

- b) *Poorly worded for unknown reasons*
“It is to work politic about land.”

Non-existent responses. The participant refuses to write any answer due to lack of sufficient comprehension and leaves comments such as “I couldn’t understand” instead.

The responses were matched against corresponding transcribed answers in order to appropriately classify them. The briefness of the video segments offsets some concern over whether participants can easily mistake the location of the answer.

In order to chart participant success in carrying out the task, a point system was allocated to each classification criteria on a rational basis. One point represented an essentially ‘Correct Response’ whereas a half point (i.e. 0.5) represented a ‘Partially-Correct Response’. Naturally, answers that were categorised as ‘Incorrect’, ‘Incomprehensible’ or ‘Non-Existent’ didn’t warrant any point value. The scoring categories are intentionally few in number in order to minimise ambiguous choices for the benefit of the author as well as appointed reliability raters. Besides, there is no need to have more categories as the important goal is simply to see if the participant could demonstrate understanding of the main point of a listening passage, even by using his/her own words. This form of top-down critical analysis on the part of the participant is deemed as more useful for this research than for them to try to just reproduce explicit details that may be overly vague, ambiguous or not make sense without the context that is normally provided through full-sentence structure.

Nevertheless, there may be a degree of dissention among educators concerning how some responses are classified. To augment a degree of external reliability into the findings so that results can be generally replicated by educators with a similar participant pool, a sample of 18 responses from 6 participants was analysed by three colleagues of the author, who teach students of similar profile (e.g. 2nd year Japanese non-majors of English from the same university). The analyses were made by these raters according to the assessment procedures outlined earlier.

Three participants were randomly chosen from the Control Group and three from the Experimental Group. One could imagine that an interrater evaluation based on teacher discretion, with a multitude of categories, may demonstrate a wide variety of interpretations of participant writing.

However, the five previously-mentioned classifications are intentionally made to be rather general and limited in number in order to make them clear and unambiguous to the rater as well as any subsequent external educator. Furthermore, the raters were able to compare participant responses with prepared transcribed answers of content taken directly from the video segments by scoring them on a point scale (i.e. 1,0.5,0).

A reasonably high level of similarity resulting from the raters' assessments and that of the author would go some way toward vindicating the 5-part classification system as a reliable means of holistically assessing learner responses. On the other hand, if there is a fairly wide discrepancy among the assessments of the raters, then this may highlight the lack of consistent and uniform standards among teachers of English as a foreign language when enforcing error correction on learner responses to listening content. Either conclusion would be useful for further research purposes.

Therefore, in order to ascertain the level of cohesion among raters, an analysis of variance in scoring (i.e. ANOVA) as well as non-linear correlation assessment between that of the author and individual raters are conducted. In order to ascertain the language level equivalence of the groups with regard to the listening-for-gist tasks, the results based on the 'College' theme are assessed according to the 5-point classification system. When distinctions, if any, are identified, they can be taken into account during the main phase of the experiment in which there is a major difference in the treatment of both groups.

Procedures for wordlist experiment

Using another theme (i.e. 'Kyoto'), the Experimental Group, unlike the Control Group, is given a sheet of paper containing a collection of key words found on the three video segments that pertain to that theme. The participants are able to use the list at their discretion. They are not limited in the time they could use to study it and may even take it home with them before doing the listening tasks. It is presumed by the instructor that some of the group members will study the list carefully with a dictionary while others may briefly scan it or even ignore it completely. There is no intention to try to force them to learn the words (if that were ethically and practically possible). The inclusion of the wordlist is meant to see whether the availability of this independent variable is effective in helping the Experimental Group fulfil the tasks more successfully than the Control Group. The wordlist is located in Appendix A. Table 1 outlines the pattern of conducting the experiment.

Table 1. Diagram of the Features of the Wordlist Experiment

Video Theme & Purpose	Control Group	Experimental Group
'College': To verify profiled groups' comparability	Without wordlist	Without wordlist
'Kyoto': To verify the impact of a wordlist	Without wordlist	With wordlist

If the Experimental Group shows a pattern that suggests the wordlist was significantly beneficial, then the view concerning the importance of vocabulary awareness in deciphering the video content among participants in relation to other factors is reinforced. However, if the results don't imply this pattern in any meaningful way, then the offer of such a wordlist (which is often demanded by students) may be a pointless tool for this holistic exercise. Naturally, for further research, the method of providing exposure to key words and/or non-lexical factors needs to be explored.

RESULTS

The Impact of Video Wordlists

This experiment set out to put the value of offering participants a presumably helpful list of key words into perspective. There are two independent groups with the same participant profile that carry out the same tasks with the same theme (i.e. College). If no significant differences in mean value are evident between the two groups on the basis of raw data in the top half of the chart in Appendix C, then finally the variable of a wordlist can be provided to the Experimental Group to see if it effectively helps its participants in providing their responses to another video theme (i.e. Kyoto). Naturally, each participant completed three responses per theme and the average score for each response reflects the scoring system of the 5-part classification system.

The unpaired t test (two-tailed p value) will examine whether the null hypothesis is true.

H_0 : The Control Group mean = The Experimental Group mean

H_1 : The Control Group mean \neq The Experimental Group mean

In accordance with the data, the Control Group mean is 1.04 and the Experimental Group mean is 1.10. Since the p value (0.69), which was calculated on the basis of an unpaired t -test result of 0.40 and degrees of freedom at 78, is greater than the significance level (0.05), the effect is not statistically significant. Therefore, the null hypothesis is accepted, meaning that there is statistically no difference between the Control and Experimental Group means.

Table 2 demonstrates a pre-test study on any differences that need to be noted between the two groups before the independent variable is integrated with the Experimental Group. The 'Participant Mean' refers to the collective points of all three segments (max. 3) of the theme, 'College'. The means could be divided by three to represent the percentage of questions that were answered correctly by both groups. When combining 'Correct' (1 point) and 'Partially-Correct' (0.5 points) results, the Control Group answered 34.7% of the questions correctly and the Experimental Group had slightly more success at 36.7%. Statistically, however, there does not appear to be any significant difference in the performance of both groups overall.

Table 2. Comparing Means and the Response Evaluations (%) of the Control Group with the Experimental Group for a Pre-Test Assessment

Theme: College	Control Group (n = 40)	Experimental Group (n = 40)
Participant Mean	1.04	1.10
Evaluations		
Correct	22.5%	16.0%
Partially-Correct	24.2%	40.0%
Incorrect	39.2%	27.5%
Incomprehensible	5.8%	9.2%
Non-Existent	8.3%	6.6%
<i>Total</i>	<i>100%</i>	<i>100%</i>

Note. Pearson Product-Moment Correlation Coefficient: 0.70

For the next experiment, the wordlist was included for the benefit of the participants of the Experimental Group. For reasons mentioned in the Literature Review, one can hypothesize that this variable will not widen the gap between the two groups in terms of their results. The raw data to be used is found in the bottom half of the chart in Appendix C with

Peter Connell

another video theme being used (i.e. Kyoto). This test involves the use of a paired *t*-test two-tailed *p* value.

H₀: The Control Group mean = The Experimental Group mean
H₁: The Control Group mean < The Experimental Group mean

On the basis of the raw data, the potential advantage given to the Experimental Group in the form of a wordlist does not appear to have taken effect. At face value, the results outlined in Table 3 show little difference between the overall mean of the Control Group (1.00) and the Experimental Group (1.18). The value of the *t*-statistic for this test is -1.02 and the two-tailed *p* value equals 0.29. The mean of the Control minus the Experimental equals -0.18 and the 95% confidence interval of this difference is from -0.50 to 0.15. The difference between 1 and 1.18 with standard deviations of 0.70 and 0.84 and sample sizes of 40 and 40, respectively, is not significant at the .05 level. Therefore, the null hypothesis is to be accepted and the alternative hypothesis is rejected. It appears to be the case that the wordlist was not particularly helpful to the Experimental Group.

Table 3. Comparing Means and Response Evaluations (%) of the Control Group with the Experimental Group Aided by a Wordlist

Theme: CO2	Control Group (n = 40)	Experimental Group (n = 40)
Participant Mean	1.00	1.18
Evaluations		
Correct	21.7%	22.5%
Partially-Correct	23.3%	32.5%
Incorrect	37.5%	31.6%
Incomprehensible	10.0%	6.7%
Non-Existent	7.5%	6.7%
<i>Total</i>	<i>100%</i>	<i>100%</i>

Note. Pearson Product-Moment Correlation Coefficient: 0.8944

There certainly is a greater discrepancy this time in favour of the Experimental Group's results based on the individual participant's mean, but not significantly enough to rule out the possibility of the element of chance. In terms of the evaluation categories displaying incorrect,

incomprehensible or non-existent responses, the 10-point gap between the two groups has remained the same. Therefore, the wordlist as a variable did not have a notable impact on the results and so the hypothesis of this research paper is accepted on the basis of the acknowledged limits of the study.

DISCUSSION

As the paired *t*-test showed an insignificant difference between the two groups in terms of their statistical means, it may be that wordlist effectiveness was offset by the impact of other error factors that are unrelated to lexical and grammatical influences. Each video segment will have a different impact on learners with various factors to consider such as ethnic accents of presenters and the speed with which they converse, program content with which participants will have differing degrees of familiarity, video picture and sound quality and so on. Individual learners may be influenced by individual factors that would require further research to uncover.

One has to also be aware of the fact that participants will study the wordlist at different levels of intensity, ranging from those who memorise the meanings of each word to others who dispense with the aid altogether. Yet even when one considers only those learners who are dedicated enough to study the list, there is still the distinct possibility that they will be unable to decipher the sounds of the words from the video.

On the basis of these complexities, an instructor may feel that a better use of time is served by just making listening transcripts accessible to students and dispensing with the idea of just giving them lists of key words. At least they will be able to learn and recognize new vocabulary as they listen to and read the passages. On the other hand, if an instructor makes the effort to show students how key words are used in sentences for particular situations before they are exposed to listening content, then there may be a greater chance for them to comprehend passages due to much of their own effort. Students may therefore feel a greater sense of achievement when they master the gist through vocabulary training, even if they miss the odd unfamiliar word that they would automatically gain through the transcripts.

The important aspect to note is that careful planning on the part of the instructor can avoid the situation in which learners are simply fed all of the words found on video content (i.e. transcripts) that don't require

Peter Connell

them to exercise their powers of critically analysing the main ideas that are expressed. In this way, the author of this paper takes issue with fellow teachers who may feel that his approach is taxing for students with little authentic listening experience behind them.

With regard to the participants of this research paper who were given a Wordlist instead of transcripts, it is apparent that they needed guidance beyond the just the opportunity to become familiar with key words used in the video segments.

IMPLICATIONS

As wordlists by themselves tended to act as superficial tools that did not seem to effectively assist participants in answering questions from the video segments, instructors may need to look at more comprehensive ways of providing learners with deeper understanding and practice of how key vocabulary can be used in order for it to be beneficial in preparation for authentic listening material.

Naturally, the findings of this research apply to Japanese learners who tend to receive much training in memorising words but little feedback in how they can be applied when expressing full answers to questions. Familiarization with words may be slightly helpful and the findings of this report do not refute that contention. Nevertheless, the common routine of students cramming the study of words before doing an oral or written test the following day without significant reinforcement has subsequently led to these words being poorly understood. It is no wonder that many of these learners lose motivation to learn a language where there is little means (e.g. time) for underpinning lexis.

Language instructors need to have the time and commitment for helping learners see how words are used in practical situations, even in the artificial contours of a classroom. This process would involve activities in which the words are presented to students in their reading, writing, speaking as well as listening activities. Once learners can gain experience both actively and passively in how key words are used, they will be not only better prepared for decoding authentic listening passages but will also be better placed to remember what they have learned over the long term.

CONCLUSION

At the very least, the experiment in this report raises doubts about the often assumed value of wordlists in relation to authentic listening content such as TV news broadcasts. What the experiment tries to reinforce is the notion that giving wordlists to learners without activities to familiarise them with how the words are effectively put into practice is on par with providing complex mathematical formulas to students who have had little exposure to reasons why they are applied.

Further research into various activities that offer pre-listening lexical training would be valuable for instructors who want their learners to be given effective assistance for adapting to the challenges of listening to the gist of British and US TV programs. Ultimately, the ideal listening material will be sufficiently demanding for even a listener with adequate knowledge of the key words contained in the audiovisual passage.

Peter Connell

NOTES

1. Fair Use Guidelines for Educational Multimedia Established by Congress of the United States (1996). Report adopted by the Subcommittee on Courts and Intellectual Property, Committee on the Judiciary, U.S. House of Representatives, on September 27, 1996, and related to Fair Use Guidelines for Educational Multimedia.

REFERENCES

- Alberding, M. (2004). Building listening skills through the basic emphasis pattern and sentence focus. *The Language Teacher*, 28(6). Retrieved May 30, 2008, from <http://www.jalt-publications.org/tlt/articles/2004/06/alberding>
- Brown, G., & Yule, G. (1983). *Teaching the spoken language*. Cambridge: Cambridge University Press.
- Daulton, F. (1998). Japanese loanword cognates and the acquisition of English vocabulary. *The Language Teacher*, 22(1). Retrieved October 4, 2007, from <http://www.jalt-publications.org/tlt/files/98/jan/daulton.html>
- Decker, M. A. (2004). Incorporating guided self-study listening into the language curriculum. *The Language Teacher*, 28(6), 5-9.
- Dunkel, P. (1986). Developing listening fluency in L2: Theoretical principles and pedagogical considerations. *The Modern Language Journal*, 70(2), 99-106.
- Freiermuth, M. (1997). L2 error correction: Criteria and techniques. *The Language Teacher*, 21(9). Retrieved April 27, 2007, from <http://www.jalt-publications.org/tlt/files/97/sep/friermuth.html>
- Fujimoto, H. (1999). The examination backwash effect on English language education in Japan. *The Japanese Learner*, 19. Retrieved April 30, 2007, from www.brookes.ac.uk/schools/education/eal/jl-archive/jl-bestof/13.pdf
- Griffie, D. (1997). Using goals and feedback to improve student performance on vocabulary homework. *The Language Teacher*, 21(7). Retrieved October 4, 2007, from <http://www.jalt-publications.org/tlt/files/97/jul/goals.html>
- Guest, M. (1998). Spoken grammar: Easing the transitions. *The Language Teacher*, 22(6). Retrieved April 20, 2007, from <http://www.jalt-publications.org/tlt/files/98/jun/guest.html>
- Hadley, G., & Mort, J. (1999). An investigation of interrater reliability in oral testing. *Research Reports of Nagaoka National College of Technology*, 35(2), 45-51. Retrieved April 30, 2007, from <http://www.nuis.ac.jp/~hadley/publication/interrater/reliability.htm>
- Hughes, A. (1989). *Testing for language teachers*. Cambridge: Cambridge University Press.
- Jacobs, H. L., Zinkgraf, S., Wormuth, D., Faye Hartfiel, V., & Hughey, J. (1981). *Testing ESL composition: A practical approach. English Composition Program*. Rowley, MA: Newbury House.

- Matsuya, Y. (2003). English teaching in Japan. *California Linguistic Notes*, 28(1). Retrieved March 21, 2007, from <http://hss.fullerton.edu/linguistics/CLN/matsuya-englishedjapan.pdf>
- McCarthy, M. (1991). *Discourse analysis for language teachers*. Cambridge: Cambridge University Press.
- Nunan, D. (1992). *Research methods in language learning*. Cambridge: Cambridge University Press.
- Obermeier, A. (2000). Listening training for TOEFL. *The Language Teacher*, 24(3). Retrieved May, 30, 2008, from <http://www.jalt-publications.org/tlt/articles/2000/03/obermeier>
- Richards, J. C. (1990). *The language teaching matrix*. New York: Cambridge University Press.
- Richards, J. C. (2003). Current trends in teaching listening and speaking. *The Language Teacher*, 27(7). Retrieved May 24, 2006, from <http://www.jalt-publications.org/tlt/articles/2003/07/richards>
- Shimo, E. (2002). Learning listening comprehension skills in English: The analysis of Japanese learners' beliefs and its implications. *The Language Teacher*, 26(10). Retrieved May 24, 2006, from <http://www.jalt-publications.org/tlt/articles/2002/10/shimo>
- South, C. (1999). The writing teacher's friend: An editing checklist for students. In *The Language Teacher*, 22(3). Retrieved April 30, 2007, from <http://www.jalt-publications.org/tlt/files/98/mar/south.html>
- Stapleton, P. (1998). Faculty word bank on the Internet. *The Language Teacher*, 22(1). Retrieved October 3, 2007, from <http://www.jalt-publications.org/tlt/files/98/jan/stapleton.html>
- Suenobu, M. (1989). *From error to intelligibility—A study of the Japanese learner's perception and production of English*. Kobe, Japan: The Institute of Economic Research, Kobe University of Commerce.
- Suenobu, M. (1995). *Communicability within errors*. Kobe, Japan: The Institute of Economic Research, Kobe University of Commerce.
- Tajima, K. (1978). The grammar-translation method: Its historical and social background. In I. Koike, H. Matsuyama, Y. Igarashi, & K. Suzuki (Eds.), *The teaching of English in Japan* (pp. 220-227). Tokyo: Eichosha Publishing.
- Weir, C. (1990). *Communicative language testing*. Hemel Hempstead, Hertfordshire, UK: Prentice Hall.
- Wu, Y. (1998). What do tests of listening comprehension test? A retrospection study of EFL test-takers performing a multiple-choice task. *Language Testing*, 15, 21-44.

CORRESPONDENCE

Peter Connell, Sir Wilfred Grenfell College, Memorial University, Canada
E-mail address: pconnell@swgc.mun.ca

APPENDIX

Appendix A. Wordlist for Experimental Group

absolutely (adv.)	individuals (n.)	pressures (n.)	threat (n.)
complex (n.)	industry (n.)	protocol (n.)	treaty (n.)
dichotomy (n.)	‘it doesn’t play’	‘put caps on’	use up (v.)
disastrous (adj.)	nuanced (adj.)	rationalize (v.)	vast (adj.)
efficient (adj.)	penalize (v.)	stance (n.)	‘walk away from’
flawed (adj.)	potentially (adv.)	suggestion (n.)	wilderness (n.)

Appendix B. Topic and Questions

Table B1. Questions Based on Each Video Segment

Title	The College Experience?
Description	A portrayal of an American college that has an unusual method of educating students.
Question #1	What is unique about Deep Springs College?
Question #2	What do most of the students do after graduation?
Question #3	What do they gain from their experience?

Table B2. Questions Based on Each Video Segment

Title	Do we need the Kyoto Protocol?
Description	Focus is placed on why the US has rejected this international agreement
Question #1	Why has the US rejected the Kyoto Protocol?
Question #2	Why are Americans not especially worried about the environment?
Question #3	What is the contradiction in US environmental standards?

Appendix C. Raw Scores for the Control and Experimental Groups

Video Theme	Control Group (n = 40)	Experimental Group (n = 40)
‘College’	0.5, 1.5, 0.5, 0, 1.5, 1, 1, 1, 0, 2, 1, 1, 0, 2, 2, 0.5, 0.5, 2, 1, 0, 2.5, 0.5, 1.5, 1.5, 0.5, 1, 1, 0, 0, 0.5, 2, 0.5, 3, 2, 1, 1, 1.5, 1.5, 0.5, 0.5	2, 0.5, 2, 0, 0.5, 2, 0.5, 1, 2, 1.5, 0.5, 0.5, 0.5, 0, 0.5, 1, 1, 1, 2, 0.5, 1.5, 2, 0.5, 1.5, 1.5, 1.5, 2, 0.5, 2, 0.5, 1.5, 1, 1.5, 1, 0.5, 1, 0.5, 1, 1, 2
‘Kyoto’	1.5, 1.5, 1, 1.5, 0, 2, 2, 1.5, 1, 1.5, 1, 1, 0.5, 0.5, 1, 1, 1, 2, 1.5, 0.5, 0.5, 0.5, 0.5, 0.5, 1, 0.5, 0.5, 0, 0.5, 1.5, 0.5, 0, 1, 3, 1.5, 1, 2.5, 0, 0, 1	1.5, 3, 1, 0.5, 2, 2.5, 0.5, 0.5, 2.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 0.5, 1, 1, 2.5, 0.5, 0.5, 2.5, 1, 1.5, 1, 0.5, 2.5, 1, 0, 0, 0.5, 2, 2.5, 2, 1, 0.5, 2, 1.5, 1.5, 0.5