The Use of Verbal Protocol Analysis in L2 Vocabulary Acquisition Patterns

Dr Beena Giridharan & Dr Chris Conlan

Curtin University

Abstract

This paper reports on a study employing Verbal Protocol Analysis (VPA) to investigate patterns of English vocabulary inferencing strategies used by second language (L2) learners between Pre-receptive and Productive processes stages for vocabulary development in pedagogical contexts. Verbal report data from 41 (n=41) tertiary-level students were gathered to deduce the processes involved in selecting inferencing strategies by L2 learners during reading. The use of verbalisations or Think Aloud Protocol provided rich data in terms of quality and content. Using a grounded approach (Glaser & Strauss, 1967; Strauss & Corbin, 1998) to data analysis, themes were teased out from the data which provide a look at how and what strategies are employed by L2 learners while reading. The study employed a mixed- method design applying a multi-dimensional approach for data gathering specifically for addressing different issues through a case study. Both concurrent and retrospective verbal protocols were gathered and were consequently transcribed, coded and evaluated for levels of understanding and strategy categorisation. These verbalisations offered insights into the cognitive learning processes of the L2 learners. The paper highlights how verbalisations inform strategy use. The study's significance lies in its ability to raise awareness in learners to be aware of their own learning with regard to vocabulary development in tertiary L2 learning contexts.

Keywords: verbal protocol analysis, tertiary L2 contexts, verbalisations and strategy use

Verbal Protocol Analysis in Reading

L2 reading research has depended upon think-aloud protocols to analyse reading processes, as reading is normally a silent, hidden process and researchers are unable to determine processes by simple observation or by a product-based assessment. Asking readers to provide verbal reports or protocols is the most direct way to access this process (Yoshida, 2008). Verbal reports include three categories of data: a) self report, where learners provide descriptions of what they do, this is generally characterised by general statements about learning behaviours; b) self-observation, which is the inspection of specific rather than generalised language behaviour soon after a learning event, and c) self-revelation, which is a "stream of consciousness disclosure" of thought processes while the information is being attended to (Cohen, 1996: 13). Verbal reporting methods that have been ascribed in educational research were used in the study because they provide data on cognitive processes. The central notion of protocol analysis is that it enables the researcher to instruct subjects to verbalise their thoughts in a way that does not modify the order of thoughts mediating the completion of a task and therefore is acceptable as valid data on thinking (Ericsson, 2002).

Study Context and Rationale

Vocabulary knowledge has been identified as the most identifiable component of the learner's ability to read (Nation & Coady, 1988).Studies in vocabulary research have revealed that L2 learners employ a series of processes for text comprehension. Understanding a text involves constructing a rational interpretation of the information presented in a text or in other words *inferencing* (Giridharan & Conlan, 2009). At a general level, inference is a cognitive process used to construct meaning. "Inference in reading comprehension is a constructive thinking process because the reader expands knowledge by proposing and evaluating competing hypothesis about the meaning of the text in an attempt to progressively refine understanding" (Davoudi,2005p.106-107). Studies on vocabulary development involving adult tertiary L2 learners are limited and research in the area of vocabulary development is also indistinct regarding the interrelationships among

various aspects of lexical competence, learning, and production processes in L2 lexical acquisition (Giridharan, 2010). Models of vocabulary acquisition in English as a second language (ESL) are scarce prompting L2 researchers to draw from first language vocabulary study models to correlate vocabulary developmental patterns. The study attempted to categorise the inferencing strategies of L2 learners and understand the patterns of vocabulary development in L2 learners of English at tertiary levels. Ultimately, the aim of the study was to formulate a theoretical model of L2 vocabulary development looking at the learner strategies applied during the inferencing process.

Research Questions Addressed

The aim of the study is to investigate patterns of inferencing strategies employed by tertiary second language (L2) learners in vocabulary acquisition, especially from '*pre-receptive vocabulary stages'* to '*productive stages'* of vocabulary acquisition.

The main objectives of the study are to:

• To study in depth the methods employed by L2 learners to infer word meanings

• To examine the transfer strategies L2 learners formulated for positive transfer during academic reading activities.

The present study introduces the concept of *pre-receptive vocabulary* as an attribute of the proficiency of the L2 learner before encountering unknown words. The belief is that *'pre-receptive vocabulary'* proficiency is a state which is prior to the stage where L2 learner receives new words for the first time. In this state the L2 learner has an established L1 schema quite unlike the individual's L2 schema and how the L2 is developed may have bearing on the foundations of language learning established earlier through the acquisition of the learner's L1. It is distinct from the belief of word familiarity which has been used to discuss the degrees of knowledge between receptive and productive vocabulary. '*Pre-receptive vocabulary'* knowledge is believed to be an essential aspect of the L2 learner that cannot be disregarded in L2 vocabulary learning.

Research Methods and Procedures

The study employed qualitative methods involving the use of verbal protocol analysis (VPA) to elicit information on the strategies used by L2 learners during reading to acquire vocabulary. VPA is also referred to as 'think alouds'. VPA was increasingly featured in this study as it distinguishes itself from other techniques that employ verbal data because the inferences are made about the cognitive processes that produced the verbalisation. VPA assisted in understanding the study participant's cognitive processes and ensured the provision of a rich data set from which to extract observations about the various cognitive and meta-cognitive strategies utilised.

VPA is a qualitative introspective technique used increasingly in second language acquisition (SLA) through which individuals' verbalisations can be perceived as accurate records of information about a particular task and involves participants 'thinking aloud' while the task is carried out (Green, 1998). Think-alouds are classified as retrospective or concurrent. In concurrent think-alouds, participants are asked to say out loud what they are thinking during the actual process of completing the task whereas, in retrospective think-alouds, participants need to recall what they were thinking while they were involved in the process of completing that task (Yoshida, 2008). In this study both concurrent and retrospective verbal protocols were utilised to investigate the inferencing patterns of adult L2 learners occurring during reading. The verbal protocol gathered was examined for description of linguistic processes and for developing themes. The verbalisations gathered in phase one of the study were recorded by the researcher and were consequently transcribed, scored and evaluated for levels of understanding. These verbalisations offered the researcher insights into the cognitive processes of the L2 learners. The study underscores developmental patterns in L2 vocabulary learning and acquisition occurring in adult tertiary learners using multiple sources of data gathered.

Participants

The 41 (n=41) L2 learners who participated in the study came from a first year Engineering undergraduate course. The researcher interacted with the participants specifically in an

Engineering communications unit which was compulsory in the course. The researcher had taught communication in engineering for five years prior to conducting the study and was actively engaged in teaching the Communications in Engineering unit during the research study period. The students in the course of study had had an essentially Malaysian background with either Mandarin (Chinese language) or *Bahasa Malaysia* (Malay language) as their first language (L1). Approximately 75% of the students at the current tertiary institution had come through the Malaysian school education system where they had learned English as a second language from standard one to standard six (primary school) up to ages 12, and from form 1 until form 5 (secondary school) until ages 17.

Purposive sampling was applied which ensured that the depth of information sought was available through the selection of participants, the site of the study and through the context in which the data was accessed. Participants were believed to have critical knowledge of the phenomenon studied, to be able to provide the etic perspective that is essential in good qualitative research, and to enhance the validity of the conclusions drawn. The data gathered from the participants provided the thick descriptions that helped develop the emergent conceptual model of L2 learning. The participants are sample representative of the population of L2 learners at a tertiary level of education and are representative of the bounded unit or group that was the focus of investigation.

Concurrent Verbalisations and Retrospective Reports

Retrospective verbal reports and concurrent verbal reports are premised within the methodology of verbal protocol analysis. In the first phase of the study, retrospective self-reports were elicited on an individual basis so that strategies suggested by one participant would not unintentionally stimulate another participant. Study participants were asked to report retrospectively on their vocabulary learning strategies following the reading of a short technical article. O'Malley & Chamot (1996) refer to the contiguity of data collection with the task about which the respondents are asked about the use of learning strategies and state that it was an essential determinant of the category of information that one can anticipate. Verbal reports are considered to provide useful information about learning strategies in

language research (Hinkel, 2005). Protocol analysis is an accurate methodology for eliciting verbal reports of thought sequences and is a valid source of data on thinking (Ericsson, 2002).

In the field of language testing, protocols are mostly expected to be gathered from subjects who are representative of the group (Green, 1998). All the participants had volunteered to participate in the study after the researcher had made an announcement to groups of first year engineering students at the institution. Firstly, each participant was given a reading text specifically related to engineering discipline containing technical vocabulary, following which they were requested to verbalise their thoughts in an audio recorder. The participants were given specific explanations on what constituted verbal protocol. Each participant provided the verbal protocol individually and was given a quiet space to read by himself/ herself while completing this activity. The researcher sat unobtrusively in another section of the room leaving the study participant to concentrate and think aloud on the task just completed. This activity was designed to recognise strategies employed in reading and comprehension of text and to understand transfer processes employed by L2 learners in developing productive vocabulary. The activity attempted to raise awareness in the L2 learner regarding their first language ability in comparison to their L2 language. In addition, it attempted to investigate the L2 learners' vocabulary retention processes. The interaction between the researcher and the participants was minimal and no attempt was made by the researcher to influence the performance of the participants other than elicit the verbalisations through written instructions and simple explanations prior to gathering the data.

The participants were informed of the proposed verbal protocol procedures and provided the choice to participate or not. Respondents who agreed to participate were provided with a letter of consent which they signed adhering to the ethical guidelines of the study. The participants were provided with a few questions on an information sheet to guide their thoughts for the verbal protocol which are shown in Table 1.

Table1 -Questions provided to guide the respondents

1. When you read articles and come across difficult words, what type of strategies do you use to understand meanings?

2. Once you have understood the words, do you use them in your own writing?

3. Can you provide some comparisons regarding your proficiency in English to your own first language?

4. What other strategies do you use to retain new words or use them as part of your vocabulary?

5. When do you consider the new words encountered as part of your mental lexicon?

"Verbal probes" are the latest additions to pure think alouds (Willis, DeMaio & Harris-Kojetin, 1999). Probe questions can be written prior to the elicitation session or can be chosen from a stock set when the elicitor judges them as appropriate, or can be created by the elicitor and are advocated by Conrad, Blair & Tracy (2000), to clarify behaviours that signal reservations or doubts. In the present study, written probe questions to guide the retrospective self-reports were provided to the participants to prompt valuable information considered crucial to ascertain the research objectives stated earlier and also to check for matches in thematic descriptions obtained subsequently from concurrent verbal protocol during the process of reading. The participants were informed of the importance of their contributions and that the disclosure of their mental processes would be informative to the research objectives, and were encouraged to verbalise their thoughts as completely as possible.

Probe questions elicited a range of relevant responses regarding vocabulary learning and inferencing patterns as seen in the transcribed responses from the participants, and also assisted to prompt the less verbal participants for more information.The validity of the verbal reports depend on how closely the method of eliciting them had adhered to the principles of the procedure and therefore it is important that appropriate instructions are used to direct the construction of verbal reports (Green, 1998). The use of guide questions as probes is considered to be un-interruptive and constructive to the data generation process. The reliability of verbal reports is dependent upon the extent to which the coding scheme accurately captures behaviour and therefore reliability of the technique is related to the validity of the coding scheme (Green, 1998).

Data Examination

Developing coding categories for the data gathered strengthened the interpretation of statements made by individual participants in that each code categorised segments of the text that referred to an inferencing strategy or vocabulary learning skill. Content analysis and analytic induction were used implicitly in the analysis of qualitative data in the study and although content was analysed qualitatively for themes and consistent patterns of meaning, a quantitative approach has been taken to content analysis through the development of theoretical units of analysis. The process necessitated the simultaneous coding of raw data and the construction of categories that capture relevant characteristics of the content. Categories and subcategories were mostly constructed through comparative methods of data analysis. Units of data that are common must be grouped; it can be something small as a word that a participant has used to describe a phenomenon (Merriam, 1998). According to Lincoln and Guba, (1995), a unit must be firstly heuristic in that it reveals information relevant to the study and stimulates the reader to think beyond the information; and also the unit should be able to stand by itself in the context in which the inquiry was carried out. Detailed theoretical units of analysis were developed from the data following appropriate coding and segmenting of the data gathered. Category construction begins with the first set of data and moves on to the next set while examining common themes or strands. This data is then compared with the next set of data. The patterns and regularities of the study become the categories or themes into which items are stored (Merriam, 1998).

Lincoln and Guba, (1995) suggest moving from concrete descriptions of data to an abstract level using concepts to describe the phenomena investigated. *Theorising* or thinking about the data helps to develop a theory that explains some aspect of the educational practice and allows the researcher to draw inferences about future activity (Merriam, 1998). Glaser and Strauss (1967) and Strauss and Corbin (1998) refer to what they call the "theoretical sensitivity" of the researcher as a useful concept with which to evaluate a researcher's skill and readiness to attempt a qualitative inquiry.

> "Theoretical sensitivity refers to a personal quality of the researcher. It indicates an awareness of the subtleties of meaning of data. ...[It] refers to the attribute of having insight, the ability to give meaning to data, the capacity to understand, and capability to separate the pertinent from that which isn't (Strauss and Corbin, 1998, p. 42).

The theory provides an explanation for the behaviour and attitudes and it may be complete with variables, constructs and hypotheses (Creswell, 2003). The study attempted to identify the emergent theories regarding inferencing strategies utilised by adult L2 tertiary learners of English through the development of a conceptual theoretical model of L2 vocabulary learning.

Second language learner strategies include both L2 learning and L2 use strategies such as retrieval, rehearsal, cover and communication strategies; the explicit goal of language learning strategies is to improve knowledge in the target language (Cohen, 1996). Some cover strategies may reflect the learner's efforts at simplification whereas in communication strategies, a learner may use a vocabulary item that he/she came across for the first time to communicate a thought, or the learner may insert the new vocabulary item into their communication to promote learning of it (Cohen, 1996).

Transcription and Coding

The data gathered were identified for patterns, consistencies, repetitions and expressions significant to the subject of the investigation. Subsequent to the verbal protocols being collected on the audio tapes, they were first transcribed and then examined to draw out codes according to a pattern or representation. The transcripts were coded independently by the investigator and inter-rater reliability was established with an experienced colleague who was a native English speaker. All protocols were coded independently by the researcher and the colleague and meetings were held to review the coded protocols for reliability checks and for discussing difficulties in coding. The verbal protocol data gathered were transcribed according to the orthographic transcription conventions recommended by Lemke (2005) who referred to thematic content as that which represents processes, activities, and relationships

and the participants in these processes, and circumstances of time, place, manner, means etc. Orthographic transcription is a verbatim record of what is said by participants which includes repetitions, pauses etc.

Appropriate symbols were developed for the transcripts through discussions with two additional raters to reflect pauses and thinking intervals. The transcripts were analysed to answer specific research questions. Green (1998) recommends the segmenting of the verbal protocols as representative of a single specific process. Green (1998) further suggests that a balance must be maintained between the researcher's desire for coding that reflects every nuance of the verbal report and the need to establish inter-rater reliability. Inter-coder reliability was established with an experienced colleague who was also a native speaker and an agreement of 98% was achieved to the allocated symbols in the transcribing process as shown in Table 2. Minor discrepancies occurred due to difficulties in recognising accents employed by the L2 learners and were resolved through discussions.

Table 2. Symbols Used for Transcribing

/=short pause

//=longer pause

[*]= filled pause

Many researchers acknowledge that transcription is an innately theoretical process dependent on the theories that the researcher upholds and influences the analysis and interpretation cycle (Chafe, 1993; Edwards, 1993; and Poland, 1995). For concurrent data analysis gathering, sixteen additional participants contributed to individual and group concurrent verbal protocol analysis. They were given a technical reading text. The reading text was selected for its use of appropriate technical vocabulary and its lexical density. In concurrent verbal protocol, the meta-cognitive processes are gathered as the task is carried out and while they occur. Each participant was requested to verbalise their meta-cognitive processes (in an audio recorder) while they read, in an effort to record strategies employed in reading comprehension and for identifying text meaning as it occurred. The procedures

followed for concurrent data gathering were similar to the retrospective self-reports and data gathering processes and adhered to the ethical guidelines outlined for the study. Research in the area of reading also focuses on the role of meta-cognition. Anderson (2002 p.1) defines meta-cognition as "thinking about thinking." The participants were provided with explanation on verbal protocol analysis to obtain a true record of the phenomenon studied. Verbal protocol analysis required subjects to give a verbal protocol (or "think" aloud) while performing the reading task. Inferencing patterns and text meaning construction methods were identified for analysis. The collection of multiple verbal protocols from a total of 41 participants, both concurrent and retrospective, provided the requisite information required for a complete analysis of inferencing strategies employed by adult L2 learners in text construction, comprehension, vocabulary learning and development. Data gathering was considered complete when the categories identified were defined, the constructs created were comprehensive, and repetitive patterns emerged from the data as a whole.
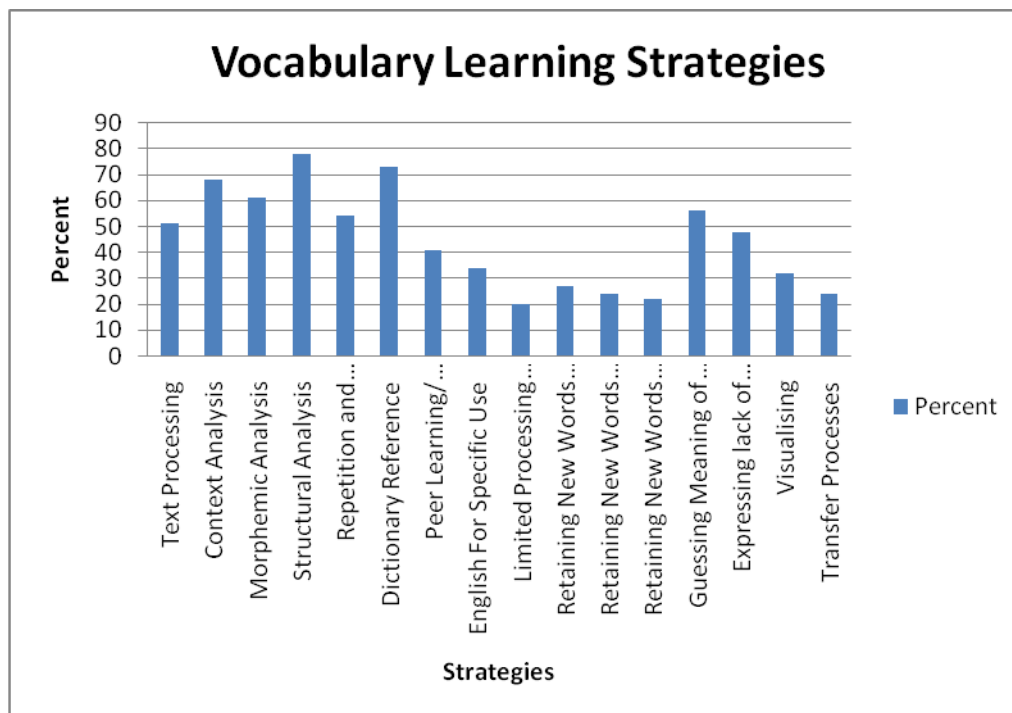
Analysis

The analyses of the transcribed verbal protocols from retrospective self reports and concurrent reports indicate that a range of strategies, such as memory, cognitive, meta-cognitive and social inferencing strategies were used by study participants during reading. Analyses of the transcribed protocol profiles were exploratory-interpretive, with the aim of categorising participants' inferencing strategies. A total of 16 major inferencing/ vocabulary learning strategies were identified from the verbal protocol analysis. Four new strategies or variants of strategies were identified from the data. Each inferencing strategy was identified as whether they correspond to memory, cognitive, meta-cognitive or a social strategy. While working on the thematic descriptions of strategies, it became clear that each participant had used inferencing strategies in unique and individual ways. Therefore, the definitions of each strategy reflect how each participant had used inferencing strategies and how these relate to descriptions of theoretical units of analysis.

Strategy Categorisation

Four new strategies were observed to have been used by the participants in the study and four strategies were observed to have been used dominantly by study participants with frequency of usage ranging between 61-78%. From the concurrent verbal protocol analysis as many as five to six strategies were documented to have been used simultaneously by participants during task performance such as reading technological texts. From the retrospective verbal protocol analysis, it was noted that participants benefited from a strong awareness regarding the inferencing strategies or vocabulary learning strategies, which augurs well for tertiary L2 learners of English.

Figure 1 shows a summary of the strategies identified in the study and the frequency of occurrence in percent.

Figure1. Summary of Identified Vocabulary Learning Strategies

It is notable from the data analysis that prior lexical knowledge played a huge role in the selection of suitable inferencing strategy for the L2 learners. Figure 1 also shows the frequency of usage of inferencing strategies utilised by study participants in the study which

were linked to numerous contextual cues available to the L2 learners, thus underscoring the significance of learning vocabulary from context. In addition to cognitive, meta-cognitive, and memory strategies employed by participants, social strategies such as asking peers and individuals in the social realm of learning required the learners to institute social communication with others, emphasising the constructivist domain of vocabulary learning. The structural analysis strategy was used most frequently by participants. This is a complex strategy that requires higher decoding skills indicative of the fact that the average of 12-13 years spent learning English as a second language tends to foster constructive vocabulary learning strategies. This pattern is evident in the study participants. Encountering unknown words is arguably an impediment to meaning construction and comprehension, however how an L2 learner negotiates meaning and selects suitable strategies to move through receptive to production processes determine the development of mental lexicons. It is interesting to note that the L2 learners in the study reiterated their convictions of the use of many vocabulary strategies previously confirmed through research studies, such as repetition and multiple exposures to words leading to productive vocabulary, which deals with incidental learning.

Discussion

Second language learner strategies include both L2 learning and L2 use strategies such as retrieval, rehearsal, cover and communication strategies. The explicit goal of language learning strategies is to improve knowledge in the target language and some cover strategies may reflect the learner's efforts at simplification whereas in communication strategies, a learner may use a vocabulary item that he/she came across for the first time to communicate a thought, or the learner may insert the new vocabulary item into their communication to promote learning of it (Cohen, 1996).

The findings from the study elucidate the various strategies and knowledge sources L2 learners refer to in order to infer meanings of unknown words encountered in the reading texts. The findings also suggest that contextual learning of word meanings may be encouraged in a tertiary L2 environment. Through contextual understanding of new words L2

learners may learn not only the syntactic and paradigmatic relations of the new words with other words but also use these associative links to form schematic patterns in their mental lexicon.

Developing semantic and thematic knowledge of content requires learners to be equipped with a range of inferencing strategies suggesting that L2 learners should be encouraged to focus attention not only on meanings or definitions of new words encountered in reading texts but also on their forms. Learners should also be made aware of the advantages of inferencing as a strategy through available contextual clues and knowledge of lexical density of texts. At university and in tertiary learning environments, L2 learners encountering low frequency words need to apply vocabulary strategies that are effectual. Hence a keen awareness of lexical and inferencing strategies are essential in L2 learners. Given that many L2 learners' achievements at a tertiary level of education depends largely in part on their ability to comprehend while reading, it is necessary to provide instruction that equips L2 learners with the lexical learning skills and inferencing strategies needed for lifelong vocabulary development.

References

Anderson, J.R. (2002). *The architecture of cognition*. Cambridge, Mass.: Harvard University Press.

Chafe, W.L. (1993). Prosodic and functional units of language. In J.A. Edwards & M.D. Lampert(Eds.), Talking data: Transcription and coding in discourse research. Hillsdale, NJ: Lawrence Erlbaum,33-43.

Cohen, A.D. (1996). Verbal reports as a Source of Insights into Second language Learner Strategies. *Applied Language Learning*, *Vol. 7*, Nos. 1 & 2,5-24.

Conrad, F., Blair, J., & Tracy, E. (2000). Verbal Reports are Data ! A Theoretical Approach to Cognitive Interviews. Retrieved 20 August 2009 http:www.fcsm.gov/99papers/conrad1.pdf.

Creswell, J. W. (2003). *Research design: Qualitative, quantitative, and mixed methods approaches.* (2nd ed.). Upper Saddle River, NJ: Pearson Education.

Davoudi, M. (2005). Inference generation skill and text comprehension. *The Reading Matrix*, *Vol. 5*, No.1, April 2005, pp. 106-123. Retrieved 20 Sept 2005: http: www. reading matrix.com/articles/davoudi/article.pdf.

Edwards, D.(1991). Categories are for talking: On the cognitive and discursive bases of categorisation. *Theory & Psychology, 1,* 515-542.

Ericsson, K.A. (2002). Protocol analysis and Verbal reports on Thinking. http://www.psy.fsu.edu/faculty/ericsson/ericsson.proto.thnk.html. Retrieved 10 June 2009.

Giridharan,B. (2010). An investigative study of English vocabulary acquisition patterns in adult L2 tertiary learners with Chinese/ Malay L1. Thesis (PhD.) Curtin University of Technology. School of Education. Faculty of Humanities. http://espace.library.curtin.edu.au.

Giridharan, B. & Conlan,C. (2009). Engaging learners in L2 vocabulary development, Teaching & Learning Forum, Curtin University of technology,17-18 July 2009.

Glaser, B.G & Strauss, A.L.(1967). The Discovery of Grounded Theory: Strategies for Qualitative Research, Chicago, Aldine Publishing Company.

Green, A. (1998). *Verbal protocol analysis in language testing research: A handbook*. Cambridge: Cambridge University Press.

Hinkel, E. (2005). *Handbook of research in second language teaching and learning.* Mahwah: Lawrence Erlbaum Associates.

Lemke, J.L. (2005). *Textual politics, discourse and social dynamics.*Taylor & Francis, London.

Merriam,S.B. (1998). *Qualitative research and case study applications in education*. London: Sage Publications.

Nation, I.S.P., & Coady, J. (1988). Vocabulary and reading. in R.Carter & M. McCarthy (Eds.). *Vocabulary and language teaching.*, London: Longman, 97-110.

O'Malley, J.M., & Chamot, A.U. (1990). *Learning strategies in second language Acquisition.* Cambridge: Cambridge University Press.

Poland, B.D.(1995). Transcription quality as an aspect of rigour in qualitative research. Qualitative Inquiry, (1), 290-310.

Strauss, A., & Corbin, J. (1998). Basics of qualitative research: Grounded theory procedures and techniques, (2nd ed.). Thousand Oaks, CA: Sage.

Willis, G.B., De Maio, T., & Harris-Kojetin, B. (1999). Is the bandwagon headed to the methodological promised land? Evaluating the validity of cognitive interviewing techniques. In Sirken, M.G., Hermann, D.J., Schechter, S., Schwartz, N., Tanur, J.M. & Tourangeau, R. (Eds). Cognition and Survey Research, Ch. 9, New York: John Wiley & Sons, 133-154.

Yoshida, M. (2008). Think-Aloud Protocols and Type of Reading Task: The Issue of Reactivity in L2 Reading Research. Selected Proceedings of the 2007 Second Language Research Forum, ed. Melissa Bowles et.al., 199-209. Somerville,M.A: Cascadilla Proceedings Project.