Experimenting the Influence of Input Modality on Involvement Load Hypothesis
Mohammad Hassanzadeh

Abstract—As far as incidental vocabulary learning is concerned, the basic contention of the Involvement Load Hypothesis (ILH) is that retention of unfamiliar words is, generally, conditional upon the degree of involvement in processing them. This study examined input modality and incidental vocabulary uptake in a task-induced setting whereby three variously loaded task types (marginal glosses, fill-in-task, and sentence-writing) were alternately assigned to one group of students at Allameh Tabataba’i University (n=21) during six classroom sessions. While one round of exposure was comprised of the audiovisual medium (TV talk shows), the second round consisted of textual materials with approximately similar subject matter (reading texts). In both conditions, however, the tasks were equivalent to one another. Taken together, the study pursued the dual objectives of establishing a litmus test for the ILH and its proposed values of ‘need’, ‘search’ and ‘evaluation’ in the first place. Secondly, it sought to bring to light the superiority issue of exposure to audiovisual input versus the written input as far as the incorporation of tasks is concerned. At the end of each treatment session, a vocabulary active recall test was administered to measure their incidental gains. Running a one-way analysis of variance revealed that the audiovisual intervention yielded higher gains than the written version even when differing tasks were included. Meanwhile, task ‘three’ (sentence-writing) turned out the most efficient in tapping learners’ active recall of the target vocabulary items. In addition to shedding light on the superiority of audiovisual input over the written input when circumstances are relatively held constant, this study for the most part, did support the underlying tenets of ILH.

Keywords—Evaluation, incidental vocabulary learning, input mode, involvement load hypothesis, need, search.

I. INTRODUCTION

It is obvious that second language (L2) learners require ample and incremental rounds of exposure to vocabulary to make sense of them and consequently internalize them, perhaps similar to what one normally experiences in L1. This means that ELT practitioners need to be wary of notions such as levels of processing and elaboration, among other terms, when teaching vocabulary. Reference [13] argues that retention of vocabulary acquired incidentally is contingent on the amount of “task-induced involvement”. By this, they simply mean deeper involvement would translate into better retention. The basic contention of the ‘Involvement Load Hypothesis’ (ILH) is that retention of unfamiliar words is, generally, conditional upon the degree of involvement in processing these words.

Incidental learning of vocabulary, among other factors, arguably requires adequate and appropriate linguistic input. The two input modes of written and oral both have their fair share of contributing to this cause. Strangely enough, while an overwhelming amount of attention is given to the written input and the relationship between reading and vocabulary, e.g. [10], [15], oral input in the form of listening activities through visuals (i.e. audiovisual input) has by and large been overlooked in the literature.

This study capitalizes on the concept of ILH, a motivational-cognitive construct which is comprised of three basic components of need, search, and evaluation [12]. Further, the audiovisual medium will be used against the reading text to examine whether vocabulary recall in the EFL context is conditional upon the task’s involvement load (i.e. the amount of need, search and evaluation it imposes) across the three tasks of marginal glosses, fill-in-blanks, and sentence-writing. What makes this re-examination unique is employing some popular TV talk shows along with reading texts resembling to them in terms of content.

II. LITERATURE REVIEW

In the domain of cognitive psychology, evidence suggests that the modality of input presentation has an effect on general human information processing [16]. Reference [16] found that the modality of presentation can influence how subjects organize information. He observed that his subjects preferred to recall information based on the modality of presentation before other criteria. When stimuli were presented in two modalities, two languages, and according to different semantic categories, subjects preferred to recall the stimuli based on the modality of presentation rather than on the language of presentation or according to semantic category.

A. Audiovisual Mode of Input

The application of video for language learning purposes within the classroom context has been summed up by [3] in their deliberation of 'technology-enhanced materials'. Speaking of non-computer resources, they commented that:

Satellite television and the more recent development of digital television both allow learners to access authentic materials from all over the world. Video clips can be presented in a single format, together with transcriptions, comprehension questions, stills and, and other items. This procedure is likely to become the norm in the near future. …These new tools offer authors and instructors a range of possibilities of presenting rich input that can be manipulated at will. [3]

Reference [4] pointed out that language found in videos could help nonnative speakers understand stress patterns.
Video allows the learner to see body rhythm and speech rhythm in second language discourse through the use of authentic language and speed of speech in various situations. Apart from viewing the language, the learner is encountered with a sociocultural environment in which the film is set.

Reaping benefits from audiovisual materials is not confined to incidental vocabulary learning. There is ample room for maneuver in other ELT-related disciplines including pragmatics and phonology for instance. Reference [1] argues that learners may optimally be exposed to pragmatic input through classroom interaction, textbook conversations, and films. A substantial number of studies [5], [8], [9], [14] suggest that classroom interaction and textbook conversations pale by comparison to the efficacy of audiovisual input. These scholars claim that authentic audiovisual input provides numerous opportunities to address all aspects of language use in a variety of contexts.

B. Involvement Load Hypothesis

The ‘need’ component in ILH is the motivational non-cognitive component of involvement and refers to whether knowledge of novel words is required to complete a task. Reference [12] further clarifies the concepts by giving examples: When a learner is reading a L2 text, s/he may encounter an unknown word necessary for understanding the text. In this case, s/he will experience a moderate need to understand the word. When the learner is writing or speaking and wants to refer to a concept but s/he does not know the word needed, s/he is also experiencing the need but at a strong degree. In other words, need is strong when it is intrinsically motivated and self-imposed by the learner.

Reference [12] refers to search and evaluation as the cognitive dimension of involvement since they involve information processing and require attention to word form and word meaning. Reference [13] holds that search is the process of looking for a L2 word for expressing a certain concept; consulting a dictionary or another authority such as the teacher. Evaluation refers to comparing of a given word with other words, a specific meaning of a word with its other meanings, or a word with its context to see whether it fits or not. A homonym like bank (river bank and a financial institution) requires the learner to make a decision as to which definition fits the specific context. A moderate evaluation entails recognizing differences between words (as in fill-in task with words provided in a list) or differences between several senses of a word in a given context. Strong evaluation requires a decision as to how additional words will combine with the new word in an original, as opposed to given, L2 sentence. Each of the three factors can be present or absent when processing a word in a natural or artificially designed task.

A task’s involvement load, then, is the combination of the presence or absence of the involvement factors of need, search, and evaluation. Reference [13] suggests that tasks with higher involvement load are deemed as more effective for word learning and retention than those with lower involvement load.

C. ILH Indices

For comparison purposes, [13] assigned tasks an involvement load index on the basis of the presence or absence of involvement factors, whereby absence of a factor is scored as 0, moderate presence of a factor as 1, and strong presence of a factor as 2. For example, a task in which learners read a text and answer comprehension questions that require knowledge of unknown words glossed in the margin would receive an involvement load index of 1 because need is moderate (imposed by the task) and search and evaluation are absent (1+ 0 + 0). In another task that requires learners to write a composition using words provided by the instructor, need is moderate (imposed by the task), search is absent, and evaluation is strong (new words are used with other words in the original text) and would attain a score of 3 (1+ 0 +2). According to the ILH, the second task is more effective than the first, because the second induces a higher involvement load.

Reference [17] examined and summarized task-induced involvement load of common language learning activities in Japanese universities which is worthy of notice for an in-depth look at the application of ILH. The depth of processing model, according to [17], stimulated the field of psychology but ILH brought more concreteness to its somewhat nebulous definition. Researchers and teachers can now employ the three components and factors in their research or teaching situations more easily than they could with the depth of processing model. Table I by [17] illustrates ILH from a more tangible perspective:

<table>
<thead>
<tr>
<th>Components</th>
<th>Degrees of the Involvement Load</th>
<th>Explanations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Need</td>
<td>index 0 (none)</td>
<td>The learner doesn’t feel the need to learn the word.</td>
</tr>
<tr>
<td></td>
<td>index 1 (moderate)</td>
<td>The learner is required to learn the word.</td>
</tr>
<tr>
<td></td>
<td>index 2 (strong)</td>
<td>The learner decides to learn the word.</td>
</tr>
<tr>
<td>Search</td>
<td>index 0 (none)</td>
<td>The learner does not look for the meanings or forms of the word.</td>
</tr>
<tr>
<td></td>
<td>index 1 (moderate)</td>
<td>The meaning of the word is found.</td>
</tr>
<tr>
<td></td>
<td>index 2 (strong)</td>
<td>The form of the word is found.</td>
</tr>
<tr>
<td></td>
<td>index 0 (none)</td>
<td>The word is not compared with other words.</td>
</tr>
<tr>
<td>Evaluation</td>
<td>index 1 (moderate)</td>
<td>The word is compared with other words in the provided context.</td>
</tr>
<tr>
<td></td>
<td>index 2 (strong)</td>
<td>The word is compared with other words in self-provided context.</td>
</tr>
</tbody>
</table>

D. Empirical Studies on ILH

ILH has been investigated through different angles. Reference [13] acknowledges that a great deal of support for ILH predates its formulation by studies not conducted to test their hypothesis. In effect, studies having a direct bearing on the hypothesis are not too many. Reference [13] conducted two parallel experiments in which their advanced Dutch and Hebrew participants (adult English learners) were formed into six intact groups. Retention of ten unfamiliar words in incidental learning setting was investigated across three task.
types (Task 1 included reading comprehension with marginal glosses, Task 2 comprehension plus filling in target words, and Task 3 composition writing with target words). The tasks had different involvement loads, i.e., various combinations of need, search and evaluation. The result indicated that Task 3 was more involving and led to better retention than Task 1 and 2, and thus lending support to ILH.

Reference [6] examined the relative importance of multiple levels of exposure and task involvement load. Students practiced target vocabulary in three conditions: one ‘fill-in-the-blank’ exercise, three ‘fill-in-the-blank’ exercises, and one ‘original-sentence-writing’ exercise. The latter had the highest involvement load. The ‘three fill-in-the-blank’ condition was significantly better (the mean score was 4.78 out of 15) than the other 2 conditions (‘one fill-in-the-blank’ was 2.18; ‘one original sentence-writing’ was 2.39). Reference [6] argued that the number of word retrievals is more important for word retention than task involvement load. Despite the overall support, he reported that his study showed word learning to be more a function of repeated exposure than involvement.

Reference [11] studied the ILH in incidental vocabulary acquisition in EFL listening. They gave three tasks to the subjects. Task A was listening comprehension questions with marginal glosses irrelevant to the questions (involvement index = 0). Task B was listening comprehension questions with marginal glosses relevant to the questions (involvement index = 1). Task C was listening comprehension questions with marginal glosses relevant to the questions and a composition writing (involvement index = 3). They found that both in immediate and delayed tests, Task C with higher involvement load produced better vocabulary retention compared to Tasks B and A.

Based on the by and large, positive outcomes analyzing ILH to date, the following research questions were floated:

1. Does ‘task-induced involvement’ significantly affect EFL learners’ incidental active word recall across audiovisual and written modes of input?

2. Is there any significant difference among the three variously loaded task types (marginal glosses, fill-in-task, and sentence-writing) across audiovisual and written input modes in terms of incidental word gains?

III. METHOD

A. Participants

The experiment included 21 Iranian undergraduate EFL students at Allameh Tabataba’i University. This stage involved only one class which was experimented dually in the form of a time-series design known as equivalent time-samples design [2]. This design was chosen for reasons of utmost subject homogeneity following [18]'s recommendation that the research designs comparing modes of input had better use participants from the same subject pool to allow for stronger claims as to whether the differences were indeed owing to modality.

B. Procedure

In order to build on the previous findings and put the ILH into experiment with respect to its benefits for incidental vocabulary acquisition, its proposed indices of ‘need’, ‘search’ and ‘evaluation’ came under scrutiny in a new light with respect to the superiority issue of input modality. The video extracts used throughout this experiment were TV ‘talk shows’ derived from the BBC’s ‘HARDTalk’ as well as the Oprah Winfrey shows. HARDTalk is a flagship BBC television program consisting of in-depth heated one-on-one interviews with distinguished people. The Oprah show is the highest-rated talk show in American television history that aired nationally for 25 seasons from 1986 to 2011.

The one and only group of learners in this phase was presented with two treatment procedures on alternate classroom sessions. Applying the construct of ILH, the teacher exposed a number of difficult lexical items (derived earlier from the extracts) through three task types assigned to the subjects during six sessions of treatment. Regarding [13]’s ILH indices (need, search and evaluation), the task types carried the following weights:

- Task A (sessions one and two) = 1 (1 + 0 + 0) i.e., moderate need + no search + no evaluation
- Task B (sessions three and four) = 2 (1 + 0 + 1) i.e., moderate need + no search + moderate evaluation
- Task C (sessions five and six) = 3 (1 + 0 + 2) i.e., moderate need + no search + strong evaluation

The class met for two back-to-back periods a week. Sessions 1 and 2 convened on the first week; Sessions 3 and 4 on week two and sessions 5 and 6 on the third week. The treatment procedure is represented in Table II.

<table>
<thead>
<tr>
<th>Session</th>
<th>Task A</th>
<th>Task B</th>
<th>Task C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Session 1</td>
<td>Vocab. checklist with L1 translation + exposure to video</td>
<td>L1 translation + exposure to video</td>
<td>Marginal glosses in L1 + reading text</td>
</tr>
<tr>
<td>Session 2</td>
<td>Fill-in-task + exposure to video</td>
<td>Fill-in-task + exposure to video</td>
<td>Fill-in-task + exposure to video</td>
</tr>
<tr>
<td>Session 3</td>
<td>Sentence-making in L2 + exposure to video</td>
<td>Sentence-making in L2 + exposure to video</td>
<td>Sentence-making in L2 + exposure to video</td>
</tr>
</tbody>
</table>

Table II: ILH TASK TYPES AND TASK INDICES ACROSS THE TWO MODES OF INPUT

---

1. Task A

Session 1 (The Audiovisual Group): Task A was arranged to have an involvement load of one. After a short warm-up discussion on plastic surgeries and its effects, students were presented with a vocabulary checklist of 16 new items along with their L1 (Farsi) equivalents transcribed earlier on the paper by the teacher from the Oprah Winfrey show. Students were asked to go over the words and their translation since they were going to appear in the extract that would follow on
the projector. A number of teacher-designed comprehension questions were also handed out to maintain the incidental component of the study. The show was an interview with a strange-looking woman who had multiple plastic surgeries. Caution was exercised as to give peripheral attention to the checklist and pretend that the core job was to answer the guiding comprehension questions. Upon watching the extract for two times, the vocabulary checklists were collected and the students were given a short time to consult before answering the comprehension questions. Soon after a random comprehension check elicitation, an active vocabulary recall test was administered on ten lexical items that had turned up on the show. With respect to scoring, each correct answer received one point, a wrong or no response received zero, and half a point was allocated to a misspelled or approximate response.

Session 2 (The Reading Group): The equivalent treatment of this task with the alternate text-as-input treatment was assigned to the very group of learners in the next period. A reading passage concerning the annals of the same woman on the talk show who was hooked on plastic surgery was given out to students. The reading, drawn up from the internet by the teacher, was in three pages and 16 difficult words chosen through an earlier pilot test appeared in boldface along with their L1 (Farsi) equivalents glossed in the margin. None of the target words nominated at this stage was identical to those used in the previous session. Learners were requested to read the passage twice before answering a number of comprehension questions. Following that, the papers were collected and the word quiz sheets were distributed to assess their active recall. Since the ‘need’ factor required in this task was moderate (induced by the teacher), both procedures were marked with ‘one’ in terms of value (see appendix for sample video and reading tests).

2. Task B

Session 3 (The Audiovisual Group): As for session three, a fill-in-task with the target words translated into L1 was handed to students to work out individually. Learners had to choose from 16 target words to go into example sentences all obtained from the exemplars in the Longman Dictionary of Contemporary English (2009). Then the 20 minute talk show recorded from BBC World News’ HARDtalk was screened twice on the projector followed by its respective recorded from BBC World News' HARDtalk was screened twice on the projector. A number of teacher-designed comprehension questions came next and once the fill-in-task was administered. This procedure bore a value of 2 based on ILH indices (moderate need, no search, and moderate evaluation).

3. Task C

Sessions 5 and 6: Finally, during sessions five and six the learners were given the vocabulary checklist along with definitions and assigned with a productive task of making new sentences with them. The sentences were then randomly read out by the students in the class and minor modifications were offered by the teacher wherever errors occurred. A talk show (Oprah) and a written text similar to what has been previously explained were distinctly brought to them during two back-to-back classroom sessions. The show was an interview with a psychologist on 'spiritual partnership after marriage’ and the reading passage was a corresponding text authored by the same psychologist. The ILH value assigned to this procedure was 3 (moderate need, no search and strong evaluation).

At the end of the treatment periods, learners’ aggregate scores on the ILH indices were summed up to put the hypothesis to the test as well as to identify the group with higher gains.

IV. RESULTS AND DISCUSSION

An ‘analysis of variance’ was used to measure our subjects’ gains. Table III illustrates the descriptive statistics for both procedures.

<table>
<thead>
<tr>
<th>Source</th>
<th>Task</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Written input</td>
<td>task A (marginal glosses)</td>
<td>3.31</td>
<td>1.33</td>
<td>21</td>
</tr>
<tr>
<td></td>
<td>task B (fill-in-task)</td>
<td>3.48</td>
<td>1.20</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>task C (sentence-writing)</td>
<td>3.85</td>
<td>1.21</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>3.54</td>
<td>1.25</td>
<td>61</td>
</tr>
<tr>
<td>Audiovisual</td>
<td>task A (marginal glosses)</td>
<td>3.95</td>
<td>.79</td>
<td>21</td>
</tr>
<tr>
<td>input</td>
<td>task B (fill-in-task)</td>
<td>3.78</td>
<td>1.37</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>task C (sentence-writing)</td>
<td>4.48</td>
<td>.86</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>4.07</td>
<td>1.06</td>
<td>61</td>
</tr>
<tr>
<td></td>
<td>task A (marginal glosses)</td>
<td>3.63</td>
<td>1.12</td>
<td>42</td>
</tr>
<tr>
<td></td>
<td>task B (fill-in-task)</td>
<td>3.63</td>
<td>1.27</td>
<td>40</td>
</tr>
<tr>
<td></td>
<td>task C (sentence-writing)</td>
<td>4.16</td>
<td>1.08</td>
<td>40</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>3.80</td>
<td>1.18</td>
<td>122</td>
</tr>
</tbody>
</table>

TABLE IV

<table>
<thead>
<tr>
<th>Source</th>
<th>Task</th>
<th>F</th>
<th>df1</th>
<th>df2</th>
<th>Sig</th>
</tr>
</thead>
<tbody>
<tr>
<td>Active vocabulary recall</td>
<td>2.135</td>
<td>5</td>
<td>116</td>
<td>.073</td>
<td></td>
</tr>
</tbody>
</table>

Levene's test of equality of variances (Table V) indicated that no significant difference among the group variances was observed (p > 0.05) and the mean scores can safely be compared.

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>F</th>
<th>Sig</th>
<th>Partial Eta Squared</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>1</td>
<td>1364.99</td>
<td>.000</td>
<td>0.92</td>
</tr>
<tr>
<td>Source</td>
<td>1</td>
<td>6.49</td>
<td>0.012</td>
<td>0.052</td>
</tr>
<tr>
<td>Task</td>
<td>2</td>
<td>2.97</td>
<td>0.055</td>
<td>0.048</td>
</tr>
<tr>
<td>Error</td>
<td>118</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected Total</td>
<td>121</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The analysis of variance (Table V) revealed that the learners’ active vocabulary recall has been significantly affected by both written and audiovisual sources of input, $F(1, 118) = 6.486, P < 0.05$. However, the higher mean scores produced in the audiovisual group ($M=4.07$) imply that as a whole, learners have accomplished more than their written exposure gains ($M=3.54$). The results at this stage once again provide further evidence to corroborate the superiority of video over the reading text when circumstances are held next to constant in tapping learners’ incidental lexical acquisition.

A second finding at this stage concerns the construct of task-induced involvement. As we know, three tasks with varying involvement indices (from 1 to 3) were designed and carried out empirically through six treatment sessions to re-examine the veracity of [13]’s proposed ILH. ANOVA results showed that incidental vocabulary learning has not been significantly affected by the type of task at 95% confidence level, $F(2, 118) = 2.967, P > 0.05$. This means that task gains in the audiovisual condition were not significantly different from the reading condition in general. Mean scores in both conditions ($video=4.07$, $reading=3.54$) indicate that audiovisual exposure accompanied by tasks yielded higher gains than the reading text. However, in the first condition (reading plus tasks) task types have yielded a linear pattern - though not significant- lending support to the theoretical foundation of ILH. In the audiovisual condition, mean scores in task 1 and 2 are in reverse order whereby task 2 is trailing behind task 1. In both conditions, the production-oriented task 3 (sentence-writing) has systematically turned out most fruitful of all while tasks 2 and 1 showed no difference from one another. Fig. 1 presents the big picture regarding learner gains.

The findings appear to be antithetical to [13]’s contention that task effect does not depend on the mode as such whether aural, oral or visual. Yet, they suggest that such predictions need to be tested and this study has followed suit. It can be firmly contended that the audiovisual mode may hold an advantage over the written mode for incidental vocabulary learning when the thematic content is kept roughly analogous.

The third task which involved sentence-writing was comprised of having learners make self-created sentences with the target words used in the video (condition A) and the reading passage (condition B). These words were practiced differently across the three task types.

In addition to shedding light on the superiority of audiovisual input over the written input when circumstances are held relatively constant, this study for the most part, did support the underlying tenets of [13]’s ILH and the proposed indices provided by them about the degree of task engagement in acquiring vocabulary. An advantage of the ILH for L2 vocabulary research is the considerable latitude and flexibility it provides for researchers in varying or manipulating the loads of the tasks based on their creativity. An interesting finding of the study was the fruitful outcome of the productive tasks. Amongst the three variously loaded tasks, the highest gains went to the third task (sentence-writing) and the video condition that enjoyed a production-requiring task.

As noted abundantly in the literature e.g. [7], language production moves learners from a primarily semantic use of language to a syntactic use where learners are urged to impose syntactic structure on their utterances. Output therefore bridges meaning and form, allowing learners to repeatedly come across instructed forms within meaningful contexts, and ultimately facilitating noticing and eventual acquisition.

### A. Limitations

We must primarily bear in mind that quantifying incidental vocabulary uptake among learners is far from straightforward and does not readily lend itself to psychometric measurement. Furthermore, the small sample size especially in the second phase may have impinged on the reliability of the findings and account for the insignificant differences obtained on task efficiencies there.

Taken together, it is safe to conclude that using appropriate audiovisual materials represents a breakthrough in vocabulary instruction and this is only one outcome that has arisen from this study. Similarly, several other aspects of second language acquisition warrant research and consideration. For instance, pronunciation practice, among others, is another category that deserves the application of video and future research may...
attend to the potential opportunities accruing from these gifts of technology.

APPENDIX

A. Sample Vocabulary Tests for Sessions 1 and 2

1. The Oprah Winfrey Talk Show Quiz (Including the Answer Key)

1) A defining example (adj.)

1. Have a second job in addition to your stomach (n.)

2) The state of being satisfied (n.)

2. A nose job (n.)

3) Sth that sticks out from a surface (n.)

3. Strong criticism (n.)

4) Use something (e.g., money) fully (v.)

4. A cosmetic operation on your face or pantyhose (n.)

5) Too keen on something (adj.)

5. Related to the distance around sth (circumferential)

6) Use this (v.)

6. A side view of sb's head (profile)

7) Suddenly and extremely (adv.)

7. A medical problem that aggravates complications (n.)

8) Modify (v.)

8. A piece of women's clothing that covers their legs (n.)

9) A piece of women's clothing that covers their legs (n.)

9. Spoiled and damaged (flawed)

10) Similarity (n.)

10. Bent or twisted (adj.)

REFERENCES


