

Does Gender Make a Difference?

A Pilot Study Using Content Based Instruction

ジェンダーの影響についての予備研究

Michael “Rube” Redfield

有名私立大学で英語必須クラスの51人の2年生たちが、学期始め（2001/4）、2つのプリントテストを受けたあと、2カ月間のコンテンツ中心教授法の単元をとった。多数選択クローズテストかリスニングクローズテストのどちらかの一方で、女子学習者（N = 29）と男子学習者（N = 22）の間には統計的な有意差はなかった。単元の終わりに、学生たちは研究データの要約に基づいた72項目のクローズテストを受けた。学力検査（X29.665）では、女子が男子よりもずっと成績がよかった。女子学生は男子よりも教室での学習に熱心で、また恐らく、より統合的に動機づけられていたのだと考えられた。

Introduction

This is the fifth in a series of studies looking at the English levels of Japanese college male and female learners. In the first study (Redfield & Zimmerman, 2001), we found no statistically significant differences between men and women English majors studying at a public college of foreign languages. The measuring instruments in this study were the structure, reading/vocabulary, and cloze sections of the MAT and the listening section of the IELTS. In a second study using undergraduate economics and business majors at a mid-ranked private university, the female students were at a significantly higher level than the men on the structure/listening CLT at the beginning of the academic year (Redfield, Bunday & Nuefer, 2001). The third study looked at those same learners after one year of English instruction. By the end of instruction, the male students had caught up to the initially higher scoring female group. A repeated ANOVA analysis showed that there were no statistically significant differences between the groups at the end of instruction, and further, that the male gain scores, but not the female gain scores, were also significant (Redfield, Bunday & Nuefer, 2002). The next study in the series (Redfield and Kotori, 2002) looked at learners enrolled in elective English conversation courses at two Japanese universities, one an elite college, the other a mid-rank college. Japanese business related majors taking elective English Conversation courses were administered a practice TOEIC exam prior to instruction. At each college, the female group posted higher means than the male group, although due

to the imbalance in gender numbers and lower sample sizes, these differences failed to show statistical significance. The results of this study indicate that in general, female non-majors are at a higher English level than male non-majors, prior to instruction. Summing up the first four studies, we find mixed results. In two cases, there were no differences in English proficiency between male and female college learners. In the other two studies, the women outperformed the men on the measuring instruments.

The present study looks at learners from an elite private university, examining academic achievement rather than general English language proficiency.

Method

Participants

The learners were recruited from two intact required sophomore English classes (N = 51), one from the Social Science Faculty (N = 37), the other from the Engineering Faculty (N = 14). The results of only those learners taking the April pre-test, the achievement test, and correctly filling out the bio-data sheet were included in the analysis. Among the fifty-one participants, there were twenty-nine females and twenty-two males.

Instruction

Identical instruction was employed in both classes, allowing us to treat the participants as a single entity. Classroom instruction was based on the fourteen part radio drama, "The Moon of India," the audio taped play that accompanies the EFL textbook East-West (Graves, 1988). The Moon of India unit takes eight ninety-minute sessions to complete. In general, two episodes are covered in each class, along with a review playing of the previous week's tape and preview of the following week's material. The test came during the eighth session. The instruction itself is tape/transcript based paired self-access learning. After listening to a new episode twice with books closed, the learners go through three question and answers sessions, with one member of dyad asking the previously prepared questions, the other answering from memory. For the second episode, the roles are reversed. After completion of the question and answer sessions, learners open the book and following along with the transcript as the tape is played once again. Finally, there is a teacher led comprehension check session. The same procedures are followed for each of the fourteen episodes.

The instructional emphasis was strictly content comprehension. Neither structure nor vocabulary were directly taught. The major idea behind the program is that second languages are

best acquired through understanding messages. If the learners, through a process of massive exposure (the tape for each episode was played a total of thirteen times over the course of three sessions, i.e., preview, self-access question and answer, and review) can come to understand the content of the radio drama, then they will, as an unconscious by-product, also acquire a significant part of the language they are being exposed to, provided of course that it is at the correct level for a majority of the learners (see Krashen, 1985).

Instruments

1). Pre-tests

There were two pre-tests; both administered during the first session in April of the 2001/2 academic year.

- a). The first measure was a twelve item multiple-choice cloze test taken from the 2001 Osaka University of Economics English Entrance Examination. It contains 137 words, scores a 62 on a Flesch Reading Ease measure, and 11 on Flesch-Kincaid Grade Level measure. Students were allowed twenty minutes to complete the cloze.
- b). The second pre-test measure was a listening cloze borrowed from Fuller (1989). The paragraph contained 198 words, and the corresponding reading statistics came out to 63.6 on the Flesch Reading Ease scale, and 8.4 on the Flesch-Kincaid Grade Level measure. Students were required to write in the correct answers. There are nineteen items to this test. The test takes approximately ten minutes.

2). Achievement test

In the *East-West Teacher' Manuel*, there are summaries of each Moon of India episode. These were pasted together to make one coherent summary of the whole radio play. Every ninth word was then deleted to make a seventy-two item cloze test used as the content-base Moon of India achievement test.

Administration.

Moon of India achievement test was administered to the students in June of 2001, one week after finishing the entire radio play. Learners were given sixty minutes to complete the seventy-two-item test. The test was scored using the one correct item only procedure.

Research question: Would gender make a difference in academic achievement?

Statistical analysis. Descriptive statistics were computed for the two pre-tests and the final achievement measure. Nonparametric procedures were used to compare the male/female results

in each case (Mann-Whitney U tests). The StatView statistical package for the Macintosh computer was used for data analysis. The alpha for statistical significance was set at .05 for this study.

Results

Pre-tests.

a). Multiple-choice cloze.

The female group (N = 29) averaged 7.828 on the multiple-choice cloze, with a standard deviation of 1.560, compared to an average of 7.682, with a standard deviation of 1.836 for the male group (N = 22). See Table One.

Table One. Multiple-choice Cloze Pre-test Descriptive Statistics.

	Count	Mean	Variance	Std. Dev.	Std. Err
female	29	7.828	2.433	1.560	.290
male	22	7.682	3.370	1.836	.391

Due to the relatively low number of participants, the nonparametric Mann-Whitney U test was used to compare the means scores of both groups. This procedure yielded a U of 315.500, with a probability of .9469, indicating that there were no statistically significant differences between groups on the pre-test multiple-choice cloze.

b). Listening cloze.

The group scores were quite similar on the listening cloze, with the female group averaging 16.069 (sd = 1.771) to the males' 15.409 (sd = 2.2856). See Table Two for descriptive statistics.

Table Two. Listening Cloze Pre-test Descriptive Statistics.

	Count	Mean	Variance	Std. Dev.	Std. Err
female	29	16.069	3.138	1.771	.329
male	22	15.409	8.158	2.856	.609

This time the Mann-Whitney procedure yielded a U of 280.500, with a probability of .4640, outside our alpha for statistical significance. In other words, there were no statistically significant differences between the two groups on this measure as well.

Achievement test.

Turning to the final achievement test, we find major differences between in performance of the gender groupings. Women learners recorded a mean score of 29.655, with a standard deviation of 9.600, while the corresponding figures for the men were 21.227 and 11.788. See Table Three for descriptive statistics.

Table Three. Achievement Test Descriptive Statistics.

	Count	Mean	Variance	Std. Dev.	Std. Err
female	29	29.655	92.163	9.600	1.783
male	22	21.227	138.946	11.788	2.513

Again using the Mann-Whitney procedure, we find a U of 190.500, with a probability of .0145, we this time find a statistical differences between the gender groups. On the post-instructional achievement cloze, the women far surpassed the men. To answer our research question, female students learned significantly more from their content-based instruction than did the men.

ANCOVA was then performed in order to check the significance of the between group differences on the achievement test, taking into consideration the non-significant differences between groups on the pre-instructional proficiency measures. See Table Four.

Table Four. Analysis of Co-variance.

	DF	Sum of Squares	Mean Square	F - Value	P - Value	Lambda	Power
sex 2	1	40.598	40.598	.509	.4796	.509	.104
classical	1	519.274	519.274	6.506	.0144	6.506	.707
listening	1	435.923	435.923	5.462	.0242	5.462	.624
sex 2 * classical	1	22.520	22.520	.282	.5980	.282	.080
sex 2 * listening	1	59.908	59.908	.751	.3911	.751	.130
classical * listening	1	389.947	389.947	4.886	.0324	4.886	.572
sex 2 * classical * listening	1	30.898	30.898	.387	.5371	.387	.091
Residual	43	3431.948	79.813				

Finally, a *post hoc* Fischer PLSD was performed. The results of this procedure (critical difference = 5.094, with a probability of .0018) confirm the original Mann-Whitney results, indicating that female achievement scores were significantly higher than the male scores.

Discussion

Although there were no statistically significant differences between the male and female scores on our two pre-test measures, the women clearly outperformed the men on our post instructional achievement test, averaging over eight points higher. Since both groups were at the same general English level prior to instruction, the only reasonable explanation for these results would be that the women were more diligent students. In the instructional approach employed, learners are basically responsible for their own learning. The bulk of the work is done in unsupervised pairs, with learners quizzing themselves in question and answer format. Looking at the results, it is obvious to this instructor/researcher that the female students as a whole take their English instruction more seriously than the men.

Given the fact that these particular learners were non-English majors, it is perhaps unsurprising, although disappointing, that the women put more effort into their classroom studies than the men. It is only speculation, but female Japanese college students seem to study English for more integrative reasons, while the men are probably more instrumentally motivated. In laymen's terms, the women seem more interested in learning English for culturally based reasons, to understand foreign movies and music, to travel, and to interact with English speakers, while the men might see English class yet another hurdle to clear for graduation. Further research into student motivation is needed in order to establish this point, however.

Conclusion

Fifty-one college sophomores taking required English classes at an elite private university, were administered two pre-tests at the beginning of the school term (April 2001), and then underwent an eight-week content-based instructional unit. There were no statistically significant differences between the female learners ($N = 29$) and the male learners ($N = 22$) on either the multiple-choice cloze test or the listening close measure. At the end of the unit, they took a seventy-two item cloze test based on a summary of the material studied. The females did much better on the achievement test ($X = 29.655$) than the males ($X = 21.227$). It was speculated that the women students were more diligent classroom earners than the males, and were perhaps more integratively motivated.

References

- Brown J. D. (1996). *Testing in language programs*. Prentice-Hall.
- Fuller, D. (1989). *Crossroads*. Macmillan Language House.
- Graves, R. (1988). *East-West*. Oxford University Press.
- Krashen, S. (1985). *The input hypothesis: Issues and implications*. Oxford: Pergamon Press.
- Kotori, Chiaki (2001). Technical Manual for the TOEIC Prep-Course Diagnostic Test. Unpublished document. University of Massachusetts/Amherst.
- Redfield, M., Bunday D., & R. Nuefer. (2001). Are There Gender Differences in English Proficiency: Looking at non English Majors. *Osaka Keidai Ronshu*, Vol. 52, No. 2, pp. 215–223.
- Redfield, M., Bunday D., & R. Nuefer. (2002). Are There Gender Differences in English Abilities among Non-Majors at the End of the Academic Year? *Osaka Keidai Ronshu*, Vol. 53, No. 4, pp. 319–324.
- Redfield, M. & C. Kotori. (2002). Are There Gender Differences in English Abilities among Non-Majors at the Beginning of the Academic Year? *Osaka Keidai Ronshu*, Vol. 53, No. 1, pp. 281–285.
- Redfield, M., & R. Zimmerman. (2001). Comparing EFL Learners in Japan: Are Men and Women the Same or Different? *Osaka Keidai Ronshu*, Vol. 51, No. 5, pp. 325–337.