UC Riverside
Issues in Applied Linguistics

Title
Examining Teacher and Student Gender Influence in Task-Prompted Oral L2 Variability

Permalink
https://escholarship.org/uc/item/6nv2g8dq

Journal
Issues in Applied Linguistics, 16(2)

ISSN
1050-4273

Authors
Rahimpour, Massoud
Yaghoubi-Notash, Massoud

Publication Date
2008-12-30

DOI
10.5070/L4162005097

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

Examining Teacher and Student Gender Influence in Task-Prompted Oral L2 Variability

Massoud Rahimpour
University of Tabriz

Massoud Yaghoubi-Notash
University of Tabriz

Variability appears to be a worthwhile strand for research in SLA arising out of the realistic, communicative approach to language learning. In accounting for variability, different frameworks have been proposed, each focusing on certain aspects of the learners. Gender is an instance of the variables that is "always present but not always apparent" (Sunderland, 2000, p. 203). In addressing the gap in literature on the relationship between gender, task and variable learner performance, this study concentrated on 20 male and 20 female university English majors' fluency, complexity and accuracy. Spoken protocols as samples of their task-prompted monologic speech addressed to the same male and female teacher were transcribed and coded for each of the three variables. Results of 2×2 (i.e. teacher gender × student gender) Repeated Measure Mixed Factorial ANOVA indicated a) overall higher fluency when addressing the female teacher, b) no significant differences in complexity in terms of neither the teacher nor the participant gender, c) females' higher accuracy regardless of the addressee, d) overall higher accuracy with the male teacher, and finally, and e) significantly higher accuracy in female participants' speech addressed to the male teacher than in any other participant-teacher pair. Implications of the study are discussed in the light of earlier findings as well as theoretical perspectives in literature.

INTRODUCTION

Meaning-focused language production on the learners' part seems to constitute the primary concern of most modern approaches to second/foreign language education (Ellis, 1994). In methodological terms, eliciting and examining samples of such productions can be of significant importance because "unless learners are given the opportunity to experience such samples they may not succeed in developing the kind of L2 proficiency needed to communicate fluently and effectively" (Ellis, 2003, p. 1). A property of these samples is variability, which concerns "how learners structure and restructure their interlanguage over time" (Ellis, 2003, p.1). In accounting for variability, various frameworks have been suggested. Within a range of "almost overwhelming" frameworks (Wolfram, 1991, p.104), some have been developed drawing upon implications from other disciplines. In what follows, dominant trends in variability accounts with their disciplinary origins are enumerated.

Issues in Applied Linguistics ISSN 1050-4273
© 2008, Regents of the University of California
Vol. 16 No. 2, 133-150

Variability Frameworks

Characterizing the variability in learners' language behavior has witnessed the rising of accumulative, succeeding paradigms over the last decades. The first paradigm to dominate emerged from and was mainly inspired by sociolinguistics. The frameworks in this category include The Labovian, Dynamic and Social Psychological paradigms. The Labovian paradigm (Labov, 1970), which has been immensely influential in shaping the earliest works on variability (Tarone, 1982, 1983), distinguishes social class, age, and gender as social factors responsible for inter-speaker variation, and stylistic factors responsible for intra-speaker variation. As clarified by Ellis (1994), styles in terms of the Labovian paradigm are

> spread along a continuum according to the amount of attention paid by the speakers to their own speech . . . . Attention serves as the mechanism through which causative social factors such as verbal task (in particular), topic, interlocutor, setting or the roles of the participants influence actual performance. (p. 122)

The Dynamic paradigm (Bailey, 1973; Bickerton, 1975), drawing on language change studies, states that varieties of language constitute a Creole continuum. Inter-speaker variation in this framework is attributed to the speakers' differential access to varieties, whereas intra-speaker variation is held to occur when speakers have access to multiple varieties. In the latter case, depending on situational factors like topic, purpose, and addressee, the speakers choose to employ one variety rather than another. Social psychological models explain variation in terms of the speakers' attitudes to in-group and out-group members (e.g. Beebe, 1988). Speech Accommodation Theory (Giles, 1971) is the most prominent social psychological framework motivating variability accounts. Three types of variation in the speech of the speakers are distinguished, namely *convergence* (when the speaker adjusts his/her speech to that of the interlocutor), *divergence* (when the speaker keeps his/her speech distinct from and dissimilar to that of the interlocutor) and *maintenance* (when the speaker makes no attempt to change his/her speech). Speech accommodation is determined by the speaker's attitude to the interlocutor(s) and can take place at any level of language use, including lexico-grammatical or discoursal levels (ibid.).

The second trend is psycholinguistic in nature and lays emphasis on the speech processing stages. The speech planning model proposed by Levelt (1989) assigns psycholinguistic sources for variability at several stages of speech production. The stages include a) a conceptualizer at which situational factors and communicative purpose shape the speaker's decision as to the variety of language, b) the formulator stage where a speech plan is made by opting for an internalized lexicon, grammar and phonological rules, c) an articulator stage, which converts the speech plan into actual speech, and d) a final stage, which enables the speaker to get feedback on his/her own speech and make phonological and grammatical adjustments (see De Bot, 1992). The speech monitoring model (Morrison and Low, 1983) resembles the

model of language production proposed by Levelt (1989); however, it further distinguishes *macro-level* (involving adjustments in terms of communicative purpose and at sentence level) and *micro-level* monitoring (lexical, syntactic and phonetic adjustments). Morrison and Low (ibid.) also propose *pre-articulatory* monitoring, which occurs before the phonetic plan is made, and *post-articulatory* monitoring, which operates on actualized speech. Skehan's cognitive approach to variability characterizes modern trends of research in which tasks have become the focus of research in their own right (See, for example, Skehan, 1998; Robinson, 2003). According to Skehan (1996, 1998), language competence is composed of formulaic lexical expressions and grammatical rules. Speakers operate the "dual processing system," which enables them to have access to both sources of knowledge. Nevertheless, depending on the communicative pressure or the accuracy demand, they have a varying dependence on lexical and grammatical processing resources (ibid.). Noting that in spontaneous production, due to the limitation of attentional sources, learners are more likely to rely on lexical processing, Skehan proposes that

> it may be possible to identify the task conditions and procedures that lead learners to place a differential emphasis on fluency, i.e. performance free of undue pauses, and false starts, complexity, i.e. the use of a wide range of grammatical structures, and accuracy, i.e. the correct use of grammatical structures. (Ellis, 2003, p. 25)

Models and frameworks reviewed so far contribute only partially to the variability accounts, and "one theory will most likely be insufficient in explaining the complexity of performance variation" (Zuengler 1989, p. 66). Therefore, the study of variability would require sociolinguistic and psycholinguistic accounts (Ellis, 1994). In being an "always present" (Sunderland, 2000, p. 203) variable in the context of language learning, gender looks appealing enough as a source of variability in SLA.

Gender, Task and L2 Variability

Although addressing gender issues in language education predates SLA, early works were almost invariably preoccupied with the so-called female superiority (see, for example, Chavez, 2001; Sunderland, 2000). Another research trend was inspired by pure gender and language studies in which male and female communicative patterns were investigated. For example, Coates' (1993) argument that females' communication is cooperative and males' is competitive and hierarchically oriented provided a good basis for classroom interaction research on male dominance in L2 situations (e.g. Spender, 1982). Concerning other aspects including communicative language use, literature on the relationship between task and gender is particularly scarce since "The TESOL profession [has] taken too long to examine gender" (Willet, 1996, p. 344).

Robinson (2001, 2003), affiliated with Skehan's cognitive perspective, identified three dimensions of tasks that cause variability in the learners' language, namely

task complexity, task difficulty, and task conditions. In this triple categorization, task conditions (as interactional factors) divides into participation and participant variables. Gender, in Robinson's (2001) terms, falls in the subcategory of participant variables. O'Sullivan (2000) could show that both males and females tended to produce more grammatically accurate forms in the presence of female interviewers, but their fluency or complexity did not vary. O'Loughlin (2002), nevertheless, in a study on the effect of the gender of the examiner in the oral interview component of IELTS, could not find any differences regarding the gender of the examiner neither quantitatively nor qualitatively. Young and Milanovich (1992) suggested that both the interviewers' and the interviewees' gender may be among the factors that bring about variations.

Considering the theoretical accounts of variability reviewed above, and also allowing for the "under-researched sites as regards gender and language learning . . . in developing countries, in Africa, Islamic countries" (Sunderland, 2000, p. 216), this paper addresses variability in Iranian context by asking the following research question:

> What is the effect of participant gender, gender of the teacher, or the interaction of both on the fluency, complexity, and accuracy of L2 learners' monologic oral L2 performance?

METHOD

Participants

Participants in the study were 20 male and 20 female sophomore and junior English majors taking Language Laboratory and Phonology courses at a private-control university in the Northwestern Iranian border town of Salmas. They were selected on the basis of a TOEFL test administration yielding two equal-sized (one all-male and the other all-female) homogeneous groups. The males' average age was 20.85, with the youngest and oldest being 19 and 25, respectively. The females' ages ranged between 20 and 26, and the average age equaled 21.65. Of the males, 12 (60%) spoke Azerbaijani, 3 (15%) spoke Persian, and the remaining 5 (25%) were the native speakers of Kurdish. With females, there were 11 (55%) Azerbaijani, 1 (5%) Persian and 8 (40%) Kurdish native speakers. (Azerbaijani and Kurdish are regional languages serving everyday communication in Iranian context. Persian is used as the official language through which almost everything, especially schooling and instruction, takes place.) The students participated in the study as part of the course assessment throughout and near the end of the autumn semester from September 2005 through February 2006 in their respective courses.

Data Collection and Procedures

Based on the participants' own rating, two topics, *City and Population* and *Love and Marriage*, were chosen from among frequently occurring speaking themes/activities contained in three dominantly used spoken English series. Although ini-

tially topic was one of the variables, in the present study it was excluded to avoid complexity in analysis.

The procedures of the study involved having the male and female participants talk once to the male and once to the female teacher as addressee on both topics. Each participant was given a two minutes of planning time before, and a five minutes of speaking time to either of the teachers. The male teacher in the experiment was the second author, and the female teacher was a departmental staff member and course lecturer with whom the participants were already acquainted. Several male and female addressees could have been used in the present study with different participants; nevertheless, the design and context of the study did not allow for different teachers. The time gap between the two task-prompted speeches addressed to the male teacher and those addressed to the female teacher was four weeks, due to local schedule considerations. Teachers as addressees merely gave the topics and initial directions, refraining from any feedback or verbal interaction during the monologues. In order to eliminate uncontrolled planning and preparation effects, arrangements were also made so that the participants who finished with their task performance could not see the ones who were waiting for their turn. The spoken protocols of the participants elicited on the four speaking events were digitally recorded.

Design

In this study, the average fluency, complexity and accuracy in the task-prompted spoken protocols of the same male vs. female participants addressed to the same male vs. female teacher were compared when talking about two topics, *love and marriage* and *city and population*. The dependent variables were fluency, complexity and accuracy (analyzed separately) with the independent variables being gender of the participant and gender of the teacher. This rendered the statistical design of the study a 2×2 (participant gender × teacher gender) Repeated Measure Mixed Factorial ANOVA (see Table 1; note that participants in rows are the same), and the methodological design is Ex Post Facto. The latter is because the whole study is a matter of manipulating research conditions so that the resulting behavior can be measured after it has occurred (see Hatch and Farhady, 1982).

Table 1. Experimental Design of the Study

	Male Teacher	Female Teacher
Male Student	N = 20	
Female Student	N = 20	

DATA ANALYSIS

The recorded data files of the participants' speech were converted to an appropriate format and analyzed with Cool Edit Pro Version 2.0, which proved especially helpful with detecting pause lengths and marking out T-unit boundaries.

Then, the recorded protocols were transcribed (see Appendix) and coded for fluency, complexity and accuracy by two independent raters.

Fluency

The ratio of meaningful words per pause (WPP) was calculated for gauging fluency. Since there are no well-defined, universally agreed-upon criteria for pauses, different local criteria are employed. 'Pause' in this study, following Crookes (1986), was operationalized as non-phonation in interclausal or intraclausal position longer than 0.60 seconds, false starts, occurrence of interword or intraword suprasegmental hesitation markers such as *mum*, *uh*, etc. (also known as filled pauses), and intraword vowel stretched longer than 0.60 seconds. In obtaining the fluency measures, repetitions, inaudible or fragmented words, unsystematic occurrence of disruption or distortion of speech by non-linguistic vocal sounds (such as coughing, sighing, etc.), as well as the words containing these occurrences word medially, were ignored. Coding for fluency did not include aspects of grammatical accuracy or mispronunciation as long as they were not meaningfully distinct. Kappa coefficients (as indices of inter-rater reliability) of the number of words and the number of pauses turned out to be 0.91 and 0.83, respectively.

Complexity

For establishing complexity of speech, different word occurrences (Types) were divided by total word occurrences (Tokens) and the result multiplied by 100. This is also known as Type-Token Ration (TTR) (see Richards, Schmidt, Platt and Schmidt, 2003). Coding for complexity disregarded sentence fragments, repeated words, incomplete clausal units, and interclausal or intraclausal interjections. The inter-rater reliability levels (indicated by kappa coefficient) were 0.90 and 0.83 for the types (i.e., the number of different words) and for the tokens (or words), respectively.

Accuracy

The general approach is to obtain the percentage of error-free T-units to the total number of T-units. T-unit is defined as "one main clause plus whatever subordinate clauses, phrases, and words happen to be attached to or embedded within it" (Menhert, 1998, p. 90). The Kappa coefficient for the error-free T-units was 0.94 and the total number of T-units was 0.81. Repetitions, fragments, and clusters of indistinct propositional link with the adjacent clausal units were left out of consideration.

RESULTS

Fluency

The 2×2 Repeated Measure Mixed Factorial ANOVA results showed that the mean word per pause (WPP) significantly varied across the speech addressed to

male vs. female teachers. However, participant gender and the interaction of the two variables did not prove significant. Results appear in Table 2.

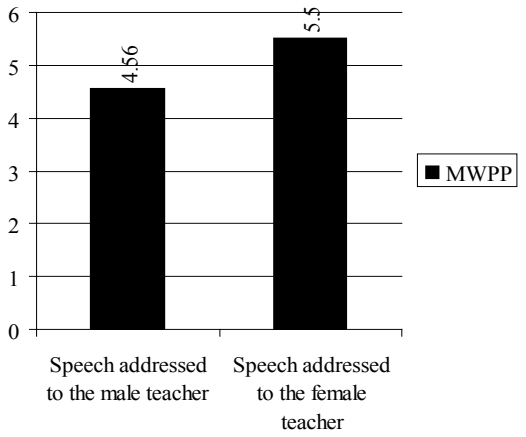
Table 2. ANOVA Table for Fluency as a Factor of Participant Gender and Teacher Gender

Tests of Within-Subject Contrasts				
Source	Type III Sum of Squares	df	Mean Square	F
Teacher gender	17.67	1	17.67	30.78**
Teacher gender × Participant gender	1.84	1	1.84	3.21 ns
Error (Teacher gender)	21.81	38	0.57	
Tests of Between-Subject Effects				
Participant gender	0.12	1	0.12	0.78 ns
Error	58.82	38	1.54	

** Significant at $p < 0.01$
ns = not significant

According to the table above, there is a significant difference in the speech addressed to the male as opposed to that addressed to the female teacher. More specifically, Figure 1 below clearly illustrates that the task-prompted speech to the female teacher contains a significantly higher number of words per pause.

Figure 1. Mean WPP (Word per Pause) in the Task-Prompted Speech Addressed to the Male vs. to the Female Teacher



Complexity

As far as the TTR (Type-Token Ratio) as the index of complexity is concerned, no significant differences could be found regarding gender of the teacher, the participant or the interaction of the two. Results obtained from ANOVA are presented in Table 3.

Table 3. ANOVA Table for Complexity as a Factor of Participant Gender and Teacher Gender

Tests of Within-Subject Contrasts				
Source	Type III Sum of Squares	df	Mean Square	F
Teacher gender	0.22	0	0.22	0.00 ns
Teacher gender × Participant gender	34.17	1	34.17	0.43 ns
Error (Teacher gender)	2045.46	38	53.82	
Tests of Between-Subject Effects				
Participant gender	178.33	1	178.33	0.18 ns
Error	3736.14	38	98.32	

** Significant at $p < 0.01$
ns = not significant

Accuracy

Regarding the percentage of error-free T-units, the statistical analysis revealed significant differences depending on gender of the participant, gender of the teacher and the interaction of the former and the latter. See Table 4 below.

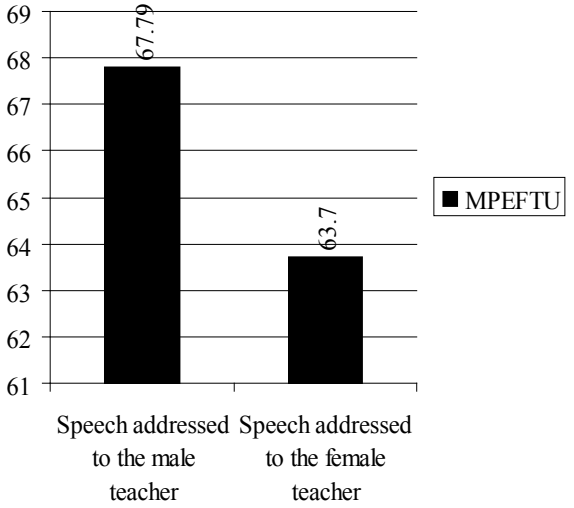
Table 4. ANOVA Table for Accuracy as a Factor of Participant Gender and Teacher Gender

Tests of Within-Subject Contrasts				
Source	Type III Sum of Squares	df	Mean Square	F
Teacher gender	518.95	1	518.95	8.65**
Teacher gender × Participant gender	358.55	1	358.55	5.97*
Error (Teacher gender)	2278.68	38		
Tests of Between-Subject Effects				
Participant gender	1479.93	1	1479.93	11.26**
Error	4992.10	38		

** Significant at $p < 0.01$
* Significant at $p < 0.05$

In line with the findings above, teacher gender appeared to influence the percentage of error-free T-units in the participants' task-prompted speech. Following Figure 2, speech to the male teacher involved more error-free T-units.

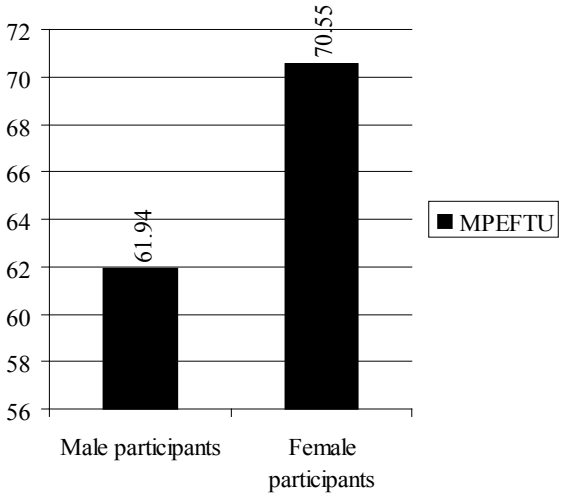
Figure 2. MPEFTU (Mean Percentage of Error-Free T-units) in the Task-Prompted Speech Addressed to the Male vs. to the Female Teacher



Another significant difference detected in terms of speech accuracy was between the male and female participants themselves. Females on average appeared

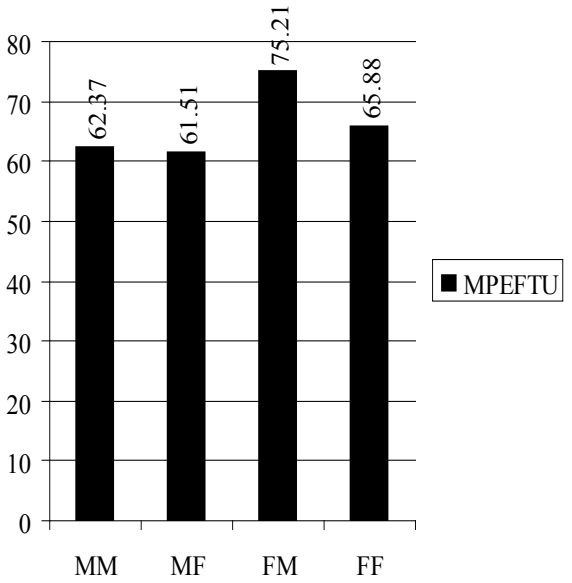
to produce a higher ratio of Error-free T-units in the speech compared to their male counterparts. Figure 3 depicts the differences graphically.

Figure 3. MPEFTU (Mean Percentage of Error-Free T-units) in the Task Prompted Speech of the Male vs. Female Participants



As far as the accuracy of the task-prompted speech is concerned, a complicated pattern emerges involving the interaction of the two variables of teacher gender and participant gender. The means of the accuracy appear in the four different columns in Figure 4 below.

Figure 4. MPEFTU (Mean Percentage of Error-Free T-units) in the Task-Prompted Speech of the Male vs. Female Participants Addressed to the Male vs. Female Teacher



MM = Male participant addressing the male teacher
MF = Male participant addressing the female teacher
FM = Female teacher addressing the male teacher
FF = Female teacher addressing the female teacher

In the case of Repeated Measure ANOVA statistical designs, the exact significance between two variables in the complex interaction of the variables can be identified through the use of a matched-pair T-test, which is the counterpart of post-hoc analyses in independent sample ANOVAs. Therefore, after applying the matched-pair T-tests, it was found that speech in the female participant-male teacher pair contained a significantly higher mean percentage of error-free T-units (MPEFTU) than any in other participant-teacher pairs in the study.

CONCLUSION AND DISCUSSION

Fluency

Results of the study indicated significant differences in terms of fluency depending on the teachers' gender. Participants' speech addressed to the female teacher was more fluent than that addressed to the male teacher. This can be attributed broadly to "interlocutor" effect in The Labovian paradigm (Labov, 1970), Dynamic paradigm (Bailey, 1973; Bickerton, 1975), and Speech Accommodation

Theory (Giles, 1971) in which variability is influenced by the addressee. Following the speech planning model by Levelt (1989), the teacher's gender can be an instance of a situational factor at the conceptualizer stage. In terms of speech monitoring (Morrison and Low, 1983), this can be associated with lexical adjustment at micro-level and pre- or post-articulatory monitoring. This finding receives general support from Robinson (2001) and Young and Milanovich (1992). However, fluent speech in the presence of a female teacher is rejected by O'Sullivan (2000) and O'Loughlin (2002), who failed to find such a variation in their studies. Higher fluency in the task-prompted speech to the female teacher may be attributed to Coates (1993), who regards women as non-hierarchically-oriented, non-debating communicators. By extension, one may argue (though based on common sense than on verifiable research) that the stereotypical image of a typical woman in the collective consciousness of the members of the society is still largely one of a soft, gentle and nurturing persona. All this might contribute to an anticipation of a tension-free and less fear-charged speech situation on the part of speakers, an instance of which was demonstrated through high fluency.

Complexity

With complexity at issue, no statistically significant differences could be found in terms of teacher or participant gender or the interaction of the two. This finding is not supported by the Labovian paradigm (Labov, 1970), Dynamic paradigm (Bailey, 1973; Bickerton, 1975), the Speech Accommodation Theory (Giles, 1971) or Levelt's (1989) formulation of situational factors, where variation is accorded with addressee influence. Conversely, the study is in line with O'Sullivan (2000) and O'Loughlin (2002), who found no differences between genders' speech. Complexity of the speech may be more a matter of linguistic resources and competence than of the speech situation and the participants involved therein. In other words, at least in this study it could be shown that the ratio of different word types occurring in actual speech may be an indication of the learners' internal competence rate rather than on who the addressee may happen to be.

Accuracy

One of the primary findings in this study on accuracy was that females were in general more accurate than males. If accuracy is equated with superiority in language, the higher accuracy of females can be attributed to the conventionally assumed female superiority in language (Chavez, 2001). Other supportive accounts are the triple categorization by Robinson (1996, 1998) and Freed and Wood (1992). Accuracy was found to vary significantly depending on the teacher's gender. Participants tended to produce a more grammatically correct L2 when addressing the male teacher. This emphasizes the interlocutor effect in prompting learner attention (Labov, 1970), situational factors in the dynamic paradigm (Bailey, 1973; Bickerton, 1975) and in Levelt's (1989) speech planning model, and participant factor as source of variability (Robinson, 2001, 2003). In particular, the findings seem to

be consistent with Morrison and Low (1983), who propose syntactic adjustments at the micro-level. Higher accuracy in addressing the male teacher is refuted by O'Sullivan (2000), who suggested the opposite, and O'Loughlin (2002), who failed to find any differences.

Higher accuracy was also found in the interaction of teacher gender and participant gender. Results indicated that females, when talking to the female teacher, were more accurate than any other participant teacher pair. In other words, females were more accurate than males irrespective of whom they talked to, and females' own speech was more grammatically correct when they talked to the male than when they addressed the female teacher. General frameworks of variability (i.e. sociolinguistic, psycholinguistic, cognitive) and claims by Freed and Wood (1996), Selinker and Douglas (1985), Robinson (2001, 2003), and Young and Milanovich (1992) apply to all three of the findings above. On the other hand, all three are rejected by O'Loughlin (2002), who dismisses any such differences. More importantly, the three findings go against those presented by O'Sullivan (2000), who claims a higher accuracy with the female interviewer. This finding can also be interpreted in the light of convergence in Speech Accommodation Theory. Assuming that the male teacher is an out-group member to the female participant (Beebe, 1988), one can suggest that females have adjusted their speech to the speech of the male as proficient authority in L2. Nevertheless, males were more inclined to maintenance than convergence or divergence with the male teacher as in-group and female as out-group members (Giles, 1971). Following Skehan (1996, 1998), female speech to the male teacher is one of the conditions that leads to communicative pressure and the requirement to be precise. With an addressee of the opposite gender, females perceived more communicative stress than did males (Ellis, 2003; Skehan, 1996, 1998), which resulted in females' dependence on the grammatical rule system, but apparently males did not experience this pressure. Another (though speculative) source of accuracy might be the fact that due to the socio-cultural factors in the Iranian context, females may have been more self-conscious in their speech addressed to the male teacher. Iran, like may other Middle Eastern countries, is for the most part a traditionally male-dominated society, and even with an increasing number of females, especially in education, males almost exclusively hold the administrative, policy and decision-making bodies.

Implications and Limitations of the Study

The findings bear several implications for SLA, including the following:

- The clearest message of this study is for task-based language teaching and learning. Gender can no longer be viewed as a static, unitary variable related to the learners only. Rather, it must be seen as an omnipresent reality inseparable from and interacting with the performances of the learners, teachers, interviewers, interviewees, peers, and groups.
- Communicative oral language testing must be adequately sensitized to the fact that the gender of the test-taker, tester, and males' and

females' attitude to the test topic may introduce bias and distort the reliability and validity of tests.

- Syllabus designers and curriculum developers must also make room for the gender of the prospective teachers, learners, and the topic within the target socio-cultural context as influential factors in determining the quantity, quality, and nature of classroom interaction.
- Language teaching practitioners need to be duly attentive to gender, which may affect or even obscure interaction in certain ways, leading to biased judgement of learner performance on the teacher's part.
- Among all other things, an intimate and at the same time an extensive understanding of dimensions of gender-related variability is called for in the realms of research, theory and practice in SLA.

Future research may shed light on various aspects of gender, which include the following:

- SLA and gender identity focusing on regional, ethnic, religious, and multicultural contexts
- Discoursal, pragmatic, sociolinguistic, and critical aspects of gender and ESL/EFL learning situation
- The negotiation of gender identity in the second/foreign language classroom and the way it clashes or aligns with norms and aspects of gendered interaction in L1

The study had several limitations. The first limitation may be argued to be the generalizability of gender influence based on 20 participants of each of the genders. This can be justified in the light of the statistically robust nature of ANOVA, which can render significant differences even with participant sizes of lower than 20. The second limitation was the time gap between the two performances of speaking to the male and to the female teachers, which may have introduced a learning effect into the learners' performance. This was due to the course restraints, and better and more generalizable results could be obtained with a 15-day time gap. Another shortcoming was that the same male and the same female teachers appeared as addressees for all 40 participants. While not losing sight of the fact that individual-specific variables may have worked their way into the performances, several addressees would have been problematic for the research in two ways. First, different addressees for the participants would not be practical, and second, using different male and female teachers would raise the criticism that speech to different teachers cannot be comparable due to different degrees of acquaintance. In anticipating the criticism, the same teachers were used to serve the addressee constancy in the research.

To conclude, the study claims to be only one step (and indeed not the only one) step in clarifying the interrelationship between gender, L2 and (task-prompted) performance. By controlling for the limitations, future research can shed more light on the role of gender in various aspects of SLA.

REFERENCES

Bailey, K. (1973). *Variation and linguistic theory*. Washington, D.C.: Center for Applied Linguistics.

Beebe, L. (1988). Five sociolinguistic approaches to second language acquisition, In Beebe, L. (Ed.), *Issues in second language acquisition: Multiple perspectives* (pp. 43-47). New York: Newbury Hourse.

Bickerton, D. (1975). *Dynamics of a Creole system*. Cambridge: Cambridge University Press.

Bischopping, K. (1993). Gender differences in conversation topic. *Sex Roles, 28,* 1-17.

Brown, G., Anderson, A., Shilcock, R., & Yule, G. (1984*). Teacher talk: Strategies for production and assessment*. Cambridge: Cambridge University Press.

Chavez, M. (2001). *Gender in the language classroom*. Boston: McGraw-Hill.

Coates, J. (1993). *Women, men, and language: A sociolinguistic account of sex differences*. New York: Longman.

Coates, J. (1997). One-at-a-time. The organization of men's talk. In S. Johnson & U.H. Meinhoff (Eds.), *Language and masculinity* (pp. 107-29). Cambridge, MA: Blackwell.

Crookes, G. (1986). *Task classification: A cross-disciplinary review.* Technical Report No. 4, The Center for Second Language Classroom Research, Social Science Research Insititute, University of Hawai'i at Manoa.

De Bot, K. (1992). A bilingual production model: Levelt's "Speaking" model adapted. *Applied Linguistics, 13,* 1-24.

Ellis, R. (1994). *The study of second language acquisition*. Oxford: Oxford University Press.

Ellis, R. (2003). *Task-based language learning and teaching*. Oxford: Oxford University Press.

Freed, A., & Wood, A. (1996). Women, men and type of talk: What makes the differences? *Language in Society, 25,* 1-26.

Gass, S., & Varonis, E. (1986). Sex differences in nonnative speaker-nonnative speaker interactions. In R. Day (Ed.), *Talking to learn: Conversation in second language acquisition* (pp. 327-351). New York: Newbury House.

Giles, H. (1971). Our reaction to accent. *New Society, 18,* 713-715.

Hatch, E., & Farhady, H. (1982). *Research design and statistics for applied linguistics*. Rowley: Newbury House.

Johnstone, B. (1993). Community and contest: Midwestern women and men creating their world in conversational storytelling. In D. Tannen (Ed.), *Gender and conversational interaction* (pp. 62-82). Oxford: Oxford University Press.

Kohonen, V. (1992). Experiential language learning: Second language learning as cooperative learner education, In D. Nunan (Ed.), *Collaborative language learning and teaching* (pp. 14-39). Cambridge: Cambridge University Press.

Kolb, D. A. (1984). *Experiential learning: Experience as the source of learning and development*. New Jersey: Prentice-Hall.

Labov, W. (1970). The study of language in its social context. *Studium Generale, 23*, 30-87.

Levelt, W. (1989). *Speaking: From intention to articulation*. Cambridge. Cambridge University Press.

Menhert, U. (1998). The effect of different lengths of time for planning on second language performance. *Studies in Second Language Acquisition, 20,* 83-106.

Morrison, D., & Low, G. (1983). Monitoring and second language learner. In J. Richard and J. Schmidt (Eds.), *Language and communication* (pp. 228-249). London: Longman.

Nunan, D. (1991). Communicative tasks and the language curriculum. *TESOL Quarterly, 25(2)*, 279-295.

O'Loughlin, K. (2002). The impact of gender in oral performance. *Language Testing, 19(2)*, 167-190.

O'Sullivan, B. (2000). Exploring gender and oral proficiency interview performance. *System, 28*, 373-386.

Richards, J.C., Schmidt, R., Platt, H., & Schmidt, M. (2003). *Longman dictionary of language teaching and applied linguistics* (3rd ed.). Essex: Longman.

Robinson, P. (2001). Task complexity, task difficulty and task production: Exploring interaction in a componential framework. *Applied Linguistics, 21*, 27-57.

Robinson, P. (2003). A cognitive hypothesis, task design, and adult task-based language learning. *Second Language Studies, 21,* 45-105.

Selinker, L., & Douglas, D. (1985). Wrestling with "context" in interlanguage studies. *Applied Linguistics, 6,* 190-204.

Skehan, P. (1996). A framework for the implementation of task-based instruction. *Applied Linguistics, 17*, 38-62.

Skehan, P. (1998). *A cognitive approach to language learning*. Oxford: Oxford University Press.

Spender, D. (1982). *Invisible women: The schooling scandal*. London: The Women's Press.

Sunderland, J. (2000). Issues of language and gender in second and foreign language education. *Language Teaching, 33*, 203-223.

Tarone, E. (1982). Systematicity and attention in interlanguage. *Language Learning, 32*, 69-82.

Tarone, E. (1983). On the variability of interlanguage systems. *Applied Linguistics, 4,* 143-63.

Willet, J. (1996). Research as a gendered practice. *TESOL Quarterly, 30,* 344-347.

Wolfram, W. (1991). Interlanguage variation: A review article. *Applied Linguistics, 12,* 102-6.

Yaghoubi-Notash, M. (2007). *Variability in L2: Exploring gender influence in task-prompted oral performance*. Unpublished PhD thesis, University of Tabriz, Tabriz, Iran.

Young, R., & Milanovich, M. (1992). Discourse variations in oral proficiency interviews. *Studies in Second Language Acquisition, 14,* 403-422.

Zuengler, J. (1989). Assessing an interactive paradigm: How accommodative should we be? In M. Eisenstein (Ed.), *The dynamic interlanguage: Empirical studies in second language variation* (pp.49-70s). New York: Plenum Press.

APPENDIX I

(0.8)	Numbers in parentheses indicate elapsed in silence in terms of a second.
(.)	A dot in parentheses indicates a small gap or pause.
::	Colons indicate prolongation of the immediately prior sound. The length of the row is indicated by the umbers in parentheses.
[ehhh]	Bracketed sounds indicate suprasegmentals.
( )	Empty parentheses indicate the transcriber's inability to hear what was said.
(word)	Parenthesized words indicate possible hearings.
(( ))	Double parentheses indicate comments interpolated by the author.
.	A period indicates a stopping fall in tone.
,	A comma indicates a continuing intonation
?	A question mark indicates rising intonation
word	Underlined parts indicate stress or higher pitch
wor--	Incomplete lexical unit

APPENDIX II

Below is the speech sample of a female student to a male teacher about *city and population*:

About population in big cities (.) oka:y [ehh] (0.8) ci—- cities include (0.9) contain many peo::ple(0.4)a large population because [uhh](0.6) because of work and living conditions. People come to big cities from countryside and o:r (0.6) small towns to work to find work. And edu—- ( )and be be educated. In this in this way they need (0.7)transport I mean bus, cars, and motorcycle and et cetra but they are polluting. Smo::ke and gas of the cars and (.)automobiles adds add smoke to the air.(0.7) In ci:ties, well, in the big cities and around them there i—- there a::re (0.6)industrial places like factories they are ((clearing throat)) they produce smoke tha::t (0.7) enters the air but also and and also they make some of the things and the chemical materials ((transfer from native language)) enter rivers, go to the go into sea and kill the fish or destroy [ehhh](1.0) the soil i::n in (1.1) the in around the rivers. In some big cities, in most ci—[uhmm] (0.85)big cities now:: (0.8)families work (.) work during the day and women work too so they don't have time to to prepare food in the homes in the home a::nd (0.65) I mean (.) that they buy food ready fast food from the shops like sandwich and everything they use everything only for that foo::d (0.8) (.) an-- so I mean they eat and (1.1) and throw away the the (0.7) (cov--) plastic cover a::nd (0.6) and or they use plastic cups plastic bags, glasses, bottles and they throw away the waste and the (0.8) environment of cities become polluted. Plastics need many years to destroy in (0.7)n—- nature and they remain on the ground for many many years and and pollution of environment happens

Massoud Rahimpour is a professor of ELT at the University of Tabriz. He holds a BA from the University of Tabriz, an M.A. degree from Oklahoma University, USA, and a PhD from the University of Queensland, Australia. He has published articles, attended conferences and supervised MA and PhD theses. He has recently been appointed Honorary Research Advisor to SLA research by the University of Queensland.

Massoud Yaghoubi-Notash is an assistant professor of ELT at the University of Tabriz. He holds a B.A from Azad University, Tabriz Branch, and an MA and a PhD from the University of Tabriz. He has presented papers at national and international conferences. He has also published co-authored papers in the proceedings of the IGALA 4 (International Gender and Language Association) Conference in Valencia, Spain.