Impact of Interlocutor Gender on Oral Proficiency Interview Interactions

Margaret Wood

Northern Arizona University

Abstract

Previous research has shown that gender impacts the language produced between two interlocutors within an interaction. This holds serious implications for speaking tests such as the Oral Proficiency Interview (OPI) which procedurally rely on face-to-face communication between two strangers. The current study investigated the impact of both interlocutor gender and gender combination of the interlocutor and the test-taker (matched or mixed) on the language elicited in the OPI. The construct was operationalized through the measurement of average length of utterance and frequency of acknowledgement tokens elicited from the test-taker. In a sample size of $N=20$, results revealed that interactions with a female interlocutor, and gender-matched interactions resulted in a longer average length of utterance and fewer acknowledgment tokens, while interactions with a male interlocutor and gender-mixed interactions resulted in the opposite: shorter average length of utterance and a higher frequency of acknowledgement tokens. This suggests that interlocutor gender in the OPI may be impacting the nature of the language elicited from the test-taker, which in turn suggests complications surrounding the validity and reliability of the OPI.

Impact of Interlocutor Gender on Oral Proficiency Interview Interactions

Background

In studies examining the effect that gender has on speech elicitation in the context of the

Oral Proficiency Interview, Kasper and Ross (2007) have acknowledged the "fundamentally

interactional quality of the interview event" (p. 2047). It is argued that various characteristics of

the interlocutor, including both conversational style and personal characteristics, have an effect

on the language elicited in the OPI setting (Kasper & Ross, 2007; Richard & Milanovic, 1992;

Seedhouse 2012). For this reason, it has been suggested that ratings achieved in OPI interactions

may not reflect just the level of the candidates' language but may also be influenced by the

nature of the data collection process. Such an extraneous variable poses an overall threat to the

construct validity and reliability of the test.

With this in mind, the current study aimed to explore the impact that gender has on data

elicitation in the OPI setting. It closely examined the data elicited in an OPI in terms of the

differences in interactions with both female and male interlocutors, and the differences found in

interactions that are gender-matched and gender-mixed.

Recent research concerning the validity and reliability of the Oral Proficiency Interview

as a test instrument suggests that individual characteristics of the interlocutor may be a variable

affecting data elicitation. According to Young and Milanovic (1992), "identities and roles of

participants, as native and non-native speakers, as experts or non-experts, and as men or women

affect the structure of the discourse," (Young & Milanovic, 1992, pp. 1-4). Conversational style

has been extensively researched in relation to gender. Merchant (2012) suggested that women

tend to exhibit less dominant styles of language – they are more polite in conversation, more

tentative, and more expressive than their counterparts. In the context of interviews, Holmes

(2013) found that women tended to be more cooperative conversationalists, adjusting their speech to that of the addressee more readily, while men seemed to be less readily responsive to the speech of others. This type of interaction in an OPI may reduce the opportunity for expansion of answers on the part of the interviewee, which is seen as a telling characteristic of oral proficiency.

It has been argued that these differences in communicative style between men and women may be rooted in social status (Holmes 2013). For example, Thorne and Henly (1975) found that women tend to interrupt less in conversation, which they suggested could be due to a lower perceived social status of the women. Zimmerman and West (1996) found that men were much less flexible in turn-taking and in the selection of topics when they were in conversation with females, which they similarly attributed to patterns of male dominance. The threat of perceived unequal social status or dominance has the potential to affect comfort level and confidence of the interviewee (Zimmerman & West, 1996, p.125). This may directly affect the language elicited in an interview due to the increased comfort – or discomfort – of the interviewee in response to the closeness of the relationship. Bilous & Krauss (1988) alternately reported that the male dominance hypothesis put forward by many researchers failed to adequately explain the conclusions generated in their research. They found that in terms of total words uttered, utterance length, interruptions, pauses, back-channeling, and laughter, the patterns were too complex to come to any decisive conclusions about the effect of gender on data elicitation (Bilous & Kruass, 1988, p.189).

Similar to Bilous and Krauss (1988), Meshkat and Javidzedeh (2006) and O'Loughlin (2002) reported that there did not appear to be any patterns falling along gender lines of the interlocutor in terms of overlaps, interruptions, or minimal responses in the context of the IELTS

OPI. In O'Loughlin's research, overlaps were almost completely absent. Interruptions were present, but their occurrence was statistically insignificant, and while minimal responses were common, they did not appear to interact with the gender variable in any definitive way. Neither O'Louglin, nor Mashkat and Javidzadeh found evidence in their results to suggest that gender had a statistically significant impact on the data (Meshkat and Javidzadeh, 2006, pp. 58 – 59).

In the context of differing sex pairings, Coates (1993) found that cross-sex pairings exhibited more frequent interruptions - particularly dominated by males. If this pattern was observed the OPI setting, it would be a serious problem in the case of gender-mixed interactions taking place between a male interviewer and a female interviewee. Alternately, Meshkat and Javidzadeh (2006) found that same-sex pairings tended to interrupt each other more frequently, which they attributed to the level of comfort that the interlocutors felt with one another (Meshkat & Javidzadeh, 2006, p. 51). Holmes (2013) found that in male to male interview contexts, men spoke to other men in a much more relaxed style, which Holmes suggested resulted from the desire to show solidarity with the interlocutor (Holmes, 2013, p.172). Similarly, Young and Milanovic (1992) found that interactions between women resulted in significantly more reactiveness by the candidate and more lengthy discussion of examiner-initiated topics (Young & Milanovic, 1992, p.14). This implies that same-sex pairings develop closer relationships with their interlocutors with more ease, which bodes well for extensive and reactive conversation in gender-matched interactions. No matter the nature of the observed pattern, research suggests that different gender-pairings of the interlocutor and interviewee may directly affect the amount and nature of the language elicited.

Kasper and Ross (2007) argued that interviewers shape interactional sequences differently from one another, and warn about the potential for this to result in construct validity

problems for the OPI. O'Sullivan and Porter (1996) suggested that in terms of measuring variables such as fluency or grammatical complexity, female interlocutors will elicit language that will ultimately receive higher ratings (p.11). The importance of eliminating variables such as these in oral proficiency exams lies in the potential for these uncontrolled variables to threaten the validity of the test. O'Sullivan and Porter stated that, "…if a feature affects test results to a significant degree, but is irrelevant to the ability being measured, it is indeed a source of measurement error which needs to be eliminated" (p. 1).

With this in mind, the present study aimed to examine the differences in language elicited from interviewees in interactions with both male and female interlocutors, and to assess the relationship between the gender-matched and gender-mixed interactions and data elicitation in the OPI. This research holds implications as to the construct validity and reliability of the OPI and provides insight into the steps forward in improving the OPI as a valid and reliable research instrument.

Research Questions

The current study sought to answer the following questions:

1) Does the gender of the interlocutor have an effect on the language elicited from the test-taker in terms of average length of utterance or acknowledgement tokens?

2) Does the gender-pairing of the interlocutor and the test-taker (matched or mixed) have an effect on the language elicited from the test-taker in terms of average length of utterance or acknowledgement tokens?

Methods

Participants

Twenty non-native speakers enrolled in undergraduate classes or ESL classes at Northern Arizona University participated as test-takers in the current study. Participants represented seven different countries: China (9), Saudi Arabia (4), Bahrain (2), Indonesia (2), Finland (1), Kuwait (1), and Mexico (1). The participant pool was comprised of 8 females and 12 males, with a mean age of 23 years old.

Four native speakers enrolled in the Ph.D. Applied Linguistics or M.A. TESL programs at Northern Arizona University participated in the current study at trained interlocutors. The participant pool comprised of 2 females and 2 males. All interlocutors had extensive formal experience in the EFL/ESL teaching context.

Materials and Data Collection Procedure

Materials consisted of a guided OPI script for the interlocutors which was adapted from material from several sources, including the "How-to" OPI YouTube channel hosted by Yijing Chen (5 December 2012), *Oral Proficiency Interview* (2017), *Oral Proficiency Interview Familiarization Manual* (2012), and other online sources from educational institutions - including Harvard University and the University of Texas (See Appendix for OPI script).

Interlocutors attended training sessions in which they were provided with a copy of the *OPI Familiarization Manual*, given verbal instruction as to how to conduct the interviews, watched authentic OPI video footage, and participated in interrater reliability exercises. They then conducted several practice interviews as a part of the pilot study, and again used these recordings to test for interrater reliability.

For the main study, recruited test-takers were matched at random with an interlocutor, and participated in a 15-25 minute, recorded Oral Proficiency Interview. Upon arrival to the interview, the test-takers were brief on the study, asked to indicate their nationality and gender identity (m/f/other/unidentified). The recordings were transcribed, timed, and analyzed for the target language features.

Language Analysis

The relationship between gender and the language elicited in the OPI was operationalized through the measurement of two variables: the amount of language elicited (measured in average length of utterance), and the frequency of acknowledgement tokens. The average length of utterance was calculated by dividing the sum of the total length of utterances by the test-taker by the total number of test-taker utterances observed. Acknowledgement tokens were counted manually and recorded.

<div align="center">

Results

</div>

Findings showed that test-takers spoke for longer on average when interacting with a female interlocutor than with a male interlocutor (Table 1), suggesting a higher degree of comfort interacting with females. Conversely, test-takers used acknowledgement tokens more frequently when interacting with a male interlocutor, suggesting a more attentive, engaged conversational style. The difference in mean proficiency level of 2 (on a 10-level scale) is not considered large enough to have a substantial effect on these data in the current study.

Table 1
Female vs. Male Interlocutor Interactions

	Average Length of Utterance (seconds)	Avg. Acknowledgement Tokens	Avg. Proficiency Level
Female Interlocutor	7.98	17	8 (Advanced Mid)
Male Interlocutor	5.99	23	6 (Intermediate High)

Results also revealed that test-takers delivered longer average utterances when in conversation with an interlocutor of the same gender (~2.5 seconds), indicating a higher level of comfort in these interactions (Table 2). As was seen in the comparison of interactions with a male versus female interlocutor, the acknowledgement tokens frequency favored the reverse; the higher number of tokens found in the gender-mixed interactions, indicating a more attentive, interactive conversation.

Table 2
Gender-Matched vs. Gender-Mixed Interactions

	Average Length of Utterance (seconds)	Avg. Acknowledgement Tokens	Avg. Proficiency Level
Gender-Matched Interaction	8.46	17	7 (Advanced Low)
Gender-Mixed Interaction	6.00	22	7 (Advanced Low)

When further broken down, data shows a substantially longer average length of utterance in female test-taker – female interlocutor interactions. Male test-taker – male interlocutor interactions were almost four seconds shorter on average and showed a nearly equal use of acknowledgement tokens (Table 3). However, it should be noted that sample size for this data set may have played a role in the variance in average length of utterance. The female - female sample pool consisted only of three total interviews, one of which was an outlier interview in which the test-taker displayed native-like proficiency.

Table 3
Gender-Matched Interactions

	Average Length of Utterance (seconds)	Avg. Acknowledgement Tokens	Avg. Proficiency Level
Female - Female Interlocutor	10.91	16	8 (Advanced Mid)
Male – Male Interlocutor	6.99	17	7 (Advanced Low)

Table 4
Gender-Mixed Interactions

	Average Length of Utterance (seconds)	Avg. Acknowledgement Tokens	Avg. Proficiency Level
Male - Female Interlocutor	6.68	18	8 (Advanced Mid)
Female – Male Interlocutor	4.98	28	5 (Intermediate Mid)

When assessing the difference between female – male interlocutor and male-female interlocutor interactions, similar patterns were found as those seen in Table 1. On average, a male speaking to a female interlocutor displayed more comfort speaking at length, (~1.5 sec. longer per utterance), while in interactions between female test-takers and a male interlocutors, acknowledgement tokens were much more frequent (Table 4).

Zimmerman & West (1996) and Thorne & Henly (1975) would have attributed the differences found in this data set to the Male Dominance Hypothesis, suggesting that a man's comfort speaking at length when interacting with a female, and a females' high level of interactivity and engagement when interacting with a male is due to the effects that perceived unequal social status has on confidence and comfort. This data also directly supports previous findings that men are less flexible in turn-taking, and women tend to be better listeners (Zimmerman & West, 1996; Thorne & Henly, 1975).

Relevance to the PIE and Second-Language Learning

Findings from the current study are relevant to the field of second-language learning in several ways. First, these findings hold strong implications as to the validity and reliability of the OPI and of speaking assessment instruments on the whole. Findings suggest that interlocutor gender in an OPI may be a variable affecting the type of language elicited from the test-taker. This, in turn, implies that amongst the tests currently in circulation that aim to measure the construct of speaking proficiency, test results may vary – meaning that creating reliable speaking assessment instruments remains a challenging task.

Second, it should be noted that differences in interactions based on gender does not end at assessment. Research has widely shown that as human beings, our interactions with others vary based on gender, and when cultural background of the interlocutor becomes an additional variable - particularly when gender norms are diverse and varied by nature - this impact becomes more pronounced. It will be important looking forward for teachers to be aware of this variable in the classroom, particularly in the ESL context, in which teachers take on the additional challenge of administering effective speaking lessons and assessment to students of a wide variety of cultural backgrounds.

References

ACTFL. (2017, September 15). List – foods and colors [Speaking novice] [Video File] Retrieved

from www.youtube.com/watch?v=lUlQKKbRZ9g&index=1&list=PLyrdUgmBeuHNy4

vId7todssm9EWquF_ZP

ACTFL. (2017, September 15). Create with language – A typical day [Speaking intermediate]

[Video File]. Retrieved from

www.youtube.com/watch?v=hyRp8sikc8w&index=2&list=PLyrdUgmBeuHNy4v

Id7todssm9EWquF_ZP

ACTFL. (2017, September 15). Create with language to have a short conversation [Speaking

intermediate] [Video File]. Retrieved from

www.youtube.com/watch?v=a0Iv_WVQC14&index=3&list=PLyrdUgmBeuHNy

4vId7todssm9EWquF_ZP

ACTFL. (2017, September 15). Ability to ask Questions [Speaking intermediate] [Video File].

Retrieved from

www.youtube.com/watch?v=FG1eMBGZsXM&list=PLyrdUgmBeuHNy4vId7tod

ssm9EWquF_ZP&index=8

ACTFL. (2017, September 15). Following up for more information [Speaking advanced] [Video

File].Retrieved from

www.youtube.com/watch?v=kGoX0EHWibY&index=11&list=PLyrdUgmBeuHN

y4vId7todssm9EWquF_ZP

ACTFL. (2017, September 15). Discussing a topic at the abstract, supporting opinion,

hypothesizing [Speaking superior] [Video File]. Retrieved from

www.youtube.com/watch?v=EzKEejU9_h0&index=6&list=PLyrdUgmBeuHNy4

vId7todssm9EWquF_ZP

ACTFL Proficiency Guidelines. (2012). ACTFL, Inc. Retrieved from www.actfl.org

ACTFL, Inc. (2012). Oral Proficiency Interview Familiarization Manual. Retrieved from

http://www.languagetesting.com/oral-proficiency-interview-opi

Bilous, F. R., & Krauss, R. M. (1988). Dominance and accommodation in the conversational

behaviours of same-and mixed-gender dyads. *Language & Communication*, *8*(34), 183-

194.

Dandonoli, P., and Henning, G. 1990. An investigation of the construct validity of the ACTFL

proficiency guidelines and oral interview procedure. *Foreign Language Annals 23*, 11-

22.

Holmes, J. (2013). *An introduction to sociolinguistics* (4th ed). New York, NY: Routledge.

Chen, Y. (2012, December 5). *Oral Proficiency Interview OPI*. Retrieved from

https://www.youtube.com/watch?v=EKA-QB9YhZ0

Coates, J. (1993). *Women, men, and language: A sociolinguistic account of gender differences in

language* (2nd ed.). London, UK and New York, NY: Longman.

Halleck, G. B. (1996). Interrater reliability of the OPI: Using academic trainee raters. *Foreign

Language Annals, 29*, 223-233.

Introduction to ACTFL Standards and the OPI (Oral Proficiency Interview). Retrieved from

https://canvas.harvard.edu/files/1081896/download?download_frd=1&verifier...

Kasper, G., & Ross, S. J. (2007). Multiple questions in oral proficiency interviews. *Journal of

Pragmatics, 39*, 2045-2070. doi:10.1016/j.pragma.2007.07.011

Liskin-Gasparro, J. *Introduction to Oral Proficiency Levels*. Retrieved from

> http://oralproficiency.coerll.utexas.edu/spanish/

Magnan, Sally Sieloff. (1986). Assessing speaking proficiency in the undergraduate curriculum:

> Data from French. *Foreign Language Annals, 19,* 429-438.

Magnan, Sally Sieloff. (1987). Rater reliability of the ACTFL oral proficiency interview. *The*

> *Canadian Modern Language Review, 43,* 267-79.

Merchant, K. (2012). *How men and women differ: Gender differences in communication styles,*

> *influence tactics, and leadership styles.* CMC Senior Thesis Paper 513.

Meshkat, M., & Javidzadeh, D. (2006). Impact of gender on discourse of oral proficiency testing

> (The Case of IELTS). *Iranian Journal of Applied Linguistics (IJAL)*, 9, 39-71.

O'Loughlin, K. (2002). The Impact of Gender in Oral Proficiency Testing. *Language Testing,*

> 19, 169-192. doi:10.1191/0265532202lt226oa

Oral Proficiency Interview Familiarization Manual. (2012). ACTFL, Inc. Retrieved from

> https://calvin.edu/dotAsset/90e68c5a-4c74-44de-8a0f-1d5c8aa7b165

O'Sullivan, B., and Porter, D. (1996). Speech style, gender and Oral Proficiency Interview

> performance. *Presented at the Annual Meeting of the Southeast Asian Ministers of*
>
> *Education Organisation*. http://files.eric.ed.gov/fulltext/ED403744.pdf

Sample Questions for Oral Proficiency Interview. (2017). Public Service Commision. The

> Government of Prince Edward Island.
>
> http://www.gov.pe.ca/psc/index.php?number=1055322&lang=E

Thorne, B. & Henley, N. (1975). Difference and dominance: An overview of language, gender,

> and society. In B. Thorne & N. Henley (Eds.), *Language and sex: Difference and*
>
> *dominance*, Rowley, MA: Newbury House Publishers, 5-42.

Young, R., & Milanovic, M. (1992). Discourse variation in oral proficiency interviews. *Studies in Second Language Acquisition*, *14*, 403-424.

Zimmermann, D. H., & West, C. (1996). Sex roles, interruptions and silences in conversation. *Amsterdam Studies in the Theory and History of Linguistic Science Series, 4,* 211-236.

Appendix

OPI Guided Questions for Interviewers

For all Interviewees:
1) Ask for participants name.
2) Can you give me a brief introduction to yourself?

Intermediate: (Create with language, ask and answer simple questions)
1) What does a typical day like?
2) What do you like to do in your free time/what do you do when you're not at work?
3) Can you tell me about your class schedule? Do you like your classes? Favorite class?
4) Can you tell me about your hometown?
5) What is your favorite book/movie/TV show/food? Can you tell me about it?
6) Can you ask me some questions about myself?

Advanced Level: (narrate, describe)
1) Tell me about the last great vacation you took.
2) Tell me about your trip over here from your home country.
3) You've studied at two different universities – how are they different or alike? (compare and contrast)
4) What was the most culturally interesting/shocking experience you have had here? How is it different than what you have experienced in your home country?
5) **Role-play:** See *Introduction to ACTFL Standards,* p.5-8

Superior Level: (give opinions, persuade, discuss, hypothesize)
1) Speak about your countries education system – what role do testing and exams play in education? Do you think it's affective?
2) What do you think is the best way to reduce human impact on climate change?
3) What can we do to encourage people in big cities to use mass transportation?
4) What role do you think pop culture plays in society? Is it positively or negatively affecting young people?
5) How often do you use social media? What is your opinion on it – do we use it too much?
6) Do attractive people have advantages that others don't?
7) Given unlimited resources, what scientific or medical problem would you investigate? Why?
8) **Role-play:** See *Introduction to ACTFL Standard,* p.5-8