LLT Journal, e-ISSN 2579-9533, p-ISSN 1410-7201, Vol. 21, Suppl, June 2018

LLT Journal: A Journal on Language and Language Teaching
http://e-journal.usd.ac.id/index.php/LLT
Sanata Dharma University, Yogyakarta, Indonesia

PAIRED ORAL TESTS: A LITERATURE REVIEW

Agustinus Hardi Prasetyo
Iowa State University, USA
hardi@iastate.edu
DOI: doi.org/10.24071/llt.2018.Suppl2110
received 2 May 2018; revised 5 June 2018; accepted 12 June 2018

Abstract
This paper reviews the studies on paired oral tests in the last ten years (2007-2017). Using the search facilities in Iowa State University's library, nine articles from some journals in the field of applied linguistics were chosen based on the inclusion criteria. Those journals are *Language Testing, Language Assessment Quarterly, Applied Linguistics*, and *Procedia – Social and Behavioral Science*. Three reasons why paired oral tests are better than interview test or individual format test are then discussed. Those are promoting and improving students' interactional competence, creating students' co-constructed discourse, and providing insights for better scale development and rater training. Paired oral tests provide opportunities for students to interact with peers in the tests, enable them to practice and improve their interactional competence. Paired oral tests also enable students to co-construct their discourse, even though there is an issue of grading the scores individually or collaboratively. The last is, more information about students' and raters' perception were gained that helps improve the rating scale and inform rater training. This paper is concluded with the call for more studies on paired oral tests to provide more insights into this complex process of creating co-constructed discourse and how to validly and reliably test both its process and product.

Keywords: pair oral test, interactional competence, co-constructed discourse

Introduction
This paper intends to review studies conducted on paired oral tests or paired speaking tests in the last ten years (2007– 2017). Paired oral tests are one type of task formats for assessing oral communication where the test takers are paired as equal speakers to have a discussion with each other (Ockey & Li, 2015). A trained rater or raters may or may not participate in the discussions. It is different from group oral tests, where more than two students are involved in the discussions, or individual format tests, where only one students who interacts with a trained rater or an assessor.

In this paper, I would like to argue that pair oral test is more beneficial than oral proficiency interview in terms of promoting and improving students' interactional competence, creating students' co-constructed discourse, and providing insights for better scale development and rater training. To conduct the review, several articles which studied paired oral test in some journals in applied linguistics field were selected. The inclusion criteria for the articles are that these

articles should be published in and after 2007, the topic is paired oral test, and those articles should be empirical research articles.

Theory

Using inclusion criteria above and the key words "paired oral test" and "paired speaking test", I searched the articles through "Quick Search" facility of the Iowa University library's online database. Besides using the quick search facility, I also used Article Indexes & Databases and e-Journal facilities to search for the articles. In fact, I also visited the websites of several journals in applied linguistics field and check the titles and the abstract of the articles which were published from the first issue of 2007 until the last issue of 2017. Nine articles were found and then selected from these following journals in applied linguistics fields: *Language Testing, Language Assessment Quarterly, Applied Linguistics,* and *Procedia – Social and Behavioral Science*. Some of the articles found were not included since they were not empirical research articles. Some of them were also not included since they discuss interview type of tests or group oral tests. In the following sections, I will discuss why paired oral tests are more superior than interview tests or individual format.

Theory Application

Students' interactional competence

All the studies reviewed in this paper mentioned that one of the advantages of paired oral test over individual format or interview type of oral test is that test takers perform better in paired oral test. Constructed within a sociocultural theory, Brooks (2009) compared the quantitative and qualitative differences in performance when the same test takers interacted with examiners and when they interacted with their peers in a test of oral proficiency. Her study was guided by these two questions: how does test-taker performance differ depending on whether the interlocutor is a tester or another student, and what are the features of interaction in the individual and paired formats? (p. 346). She claimed that test takers who participated in paired format scored better than when they participated in the individual format (when they interacted with an examiner). Moreover, the qualitative analyses of the interactional discourse elicited during paired oral tests showed that more interaction, negotiation of meaning, and complex output were produced. Test-takers employed more features of interaction (17 features) in paired test, while in the individual format the test takers employed 10 features of interaction. Moreover, from the Conversation Analysis conducted by the researcher, it was found out that the interaction was more asymmetrical in nature, similar to that in an interview. This result supported the findings of previous studies that pair format is better than interview or individual format in terms of students' performances.

A study conducted by Laborda, Juan, and Bakieva (2015) also yielded similar result. They conducted a study to test the construct of the new Spanish University Entrance Examination (PAU) where an experimental paired oral tests format was conducted with potential participants of PAU. Laborda et al. concluded that co-construction of output resulted from paired oral tests format

supported the development of students' interactional competence and improved individual student's performance. They further claimed that in paired oral tests, test takers tended to support their peers' responses. This might have a significant effect on the students' performances. Moreover, the atmosphere was relaxing since it was their friends they were addressing. The test takers tended to speak better and more so the length of their discourse also increased.

Galaczi (2008) conducted a study that investigated the relationship between the score of interactional competence that the test takers received in their paired oral tests and their pattern of interaction in their co-constructed discourse in paired oral tests. She found out that there were three patterns of interactions in the discourse: collaborative, parallel, and asymmetric. In collaborative interaction, the test takers were mutually and equally engaged in the interaction. It means that they were actively engaged in the co-construction of discourse. The second is parallel interaction, where the students were not mutually nor equally engaged in the interaction. It is like "solo vs. solo" interaction. In the third interaction, asymmetric interaction, one of the participants was dominant, while the other was passive. She also found that there was a significant correlation between the students' score in their interactional competence and their pattern of interactions. Test takers who were mutually and equally engaged, who were actively co-constructing their discourse were proven to have higher scores in their interactional competence than those test takers who had parallel or asymmetric interaction. In another study, May (2009) also showed clearly that paired oral test could elicit features of interactional competence, including conversation management skills, that cannot be captured or even do not exist in interview or individual oral type of test. Those features of interactional competence can be best elicited through tasks involving test takers' interaction.

All these studies then show that paired oral test helps promote and improve test takers' interactional competence. In the following section, I will discuss the next feature of paired oral test that makes it better than individual format test: the creation of students' co-constructed discourse.

The creation of students' co-constructed discourse

The term interactional competence was first coined by Kramsch (1986) who argued that since the interactional discourse is co-constructed by participants involved in it, the responsibility for that discourse cannot be assigned to just one participant involved in that discourse construction. Or in a paired oral test setting, the score of interactional competence cannot be assigned to just one test taker, but it must be shared equally by all the test takers involved. This paired oral test setting then creates an opportunity as well as a challenge. On one hand, paired oral tests enable the creation of rich and more authentic discourse, which resulted from the process of negotiating meaning and not just information transfer. On the other hand, it raises the issue of validity and fairness. How valid is the score of interactional competence awarded to the test takers? How fair is the score awarded? What if one participant of the paired oral tests was low or weak in terms of their interactional competence or linguistic ability?

Ducasse and Brown (2009) and May (2009) conducted a study about these issues viewed from the raters' perspectives. Ducasse and Brown (2009) reported

the findings of verbal protocols of teacher-raters who observed the paired oral test discourses. These verbal protocols gave insights on what raters were focusing on when rating paired oral examinees. The focus of their study was therefore on the construct of interaction. The findings reveal that the raters observed and identified in the students' co-constructed discourse in paired oral tests three main categories of interactional features: non-verbal interpersonal communication (which has two subcategories: gaze and body language), interactive listening (with two subcategories: supportive listening and comprehension), and interactional management (with also two subcategories: horizontal and vertical management). The definition of the construct of effective interaction between examinees in paired oral tests should therefore take into account these interactional features, since those are what the raters are considering when rating the examinees. Also, those interactional features should be considered in the development of rating scales. The results of their study then provide insights on how to create more valid and fair test scale to assess students' interactional competence depicted through the creation of co-constructed discourse.

A similar study conducted by May (2009) who also argued that since the interaction in a paired oral or speaking test is intrinsically co-constructed in nature, giving shared scores for the test-takers' interactional competence is one way of acknowledging it. Her study showed that it is difficult for raters to assign scores to test takers, especially when their nature of interaction is asymmetrical, where one participant is dominant and the other is passive. She suggested that in order for the paired oral tests to be fair and valid, each test taker still should still receive a separate score for Accuracy, Fluency, and Range (p. 417)

If those two previous studies discussed the students' co-constructed nature of paired oral tests from the raters' perspectives, Bennett (2012), Davis (2009), and Lazaraton and Davis (2008) discussed it from test takers' perspectives. Lazaraton and Davis (2008) argued that test takers bring their language proficiency identity (LPID) to the test tasks, and this identity is fluid. It means the test takers' identity changes, depending on who their interlocutor is. In their study, using the notion of "positioning", they found that the test takers' LPID can manifest in the talk by "do being proficient", "do being interactive", "do being supportive", and "do being assertive". Do being proficient and do being interactive mean that the overall proficiency that the test takers show synergistically and collaboratively positions them as competent English users, therefore they deserve high scores on the paired oral test. Do being supportive and do being assertive take place in a talk involving a more proficient speaker with a weaker one. They also deserve high scores with those identities. Based on the results of their study, Lazaraton and Davis recommended that the test takers should be tested twice with different partners to find out what their true LPID is.

Davis (2009) in his study found that the proficiency level of test takers' interlocutor or partner in a pair oral test has no effect on the test takers' performance. Higher-proficiency test takers were generally not harmed by interacting with a lower-level test taker. However, lower-level student did not greatly benefit from working with a higher-level peer either, at least in terms of score. He also found that in his study, most of the conversations produced

collaborative interaction. This supported Galaczi's (2008) study, that there is a global pattern of interactions in the test takers' co-constructed discourse, namely collaborative interaction (where the test takers are mutually and equally engaged), parallel interaction (where both speakers are equal, initiated and developed topics, but not mutual, which means they are not engaged with each other's ideas), and asymmetric interaction (where one speaker is passive and the other is dominant). Bennett (2012) also found that interlocutor's linguistic ability has little or no influence on the test taker's performance. In fact, based on the post-test questionnaire, the test takers felt satisfied with the pairing.

The last benefit of paired oral tests that I would like to discuss is the insights and understanding of better scale development and rater training gained from studies conducted on paired oral tests.

Insights for scale development and rater training

Galaczi (2014) conducted a study on interactional competence within varying proficiency levels, in this case CEFR proficiency level. The data of her study were 41 average pairs selected from the 84 video-taped test taker performances on the test taker interaction task at CEFR levels B1 to C2 or four proficiency levels. The term average here refers to test takers who had a mark 3-4 band (from a 1-5 band scale) on the Cambridge English Interactive Communication scale. She employed a mixed-methods approach (Creswell, 2014), combining a contrastive analysis technique and quantitative coding of the data. The research question of her study was "what features of interactional competence in test-taker discourse are salient at different oral proficiency levels?". The results of contrastive analysis showed that several interactional features distinguish proficiency levels. The test takers in the four proficiency levels were engaging in the three key interactional features: topic development, listener support, and turn-taking management. This study then gave insights to the conceptualization of the Interactional Competence construct by providing useful descriptive interactional features which could
supplement the already available Interactional Competence scales and descriptors.

Other studies reviewed in this article also argued that their studies will give insights into the development of scale and rater training. May's (2009) study is claimed to provide insights into raters since it investigated raters' perceptions on whether they considered separable the individual contribution to interactional patterns in paired oral tests. May claimed that her study will provide insights into the development of rating scales which can capture the complexities of interactional competence in a paired oral test, and the training of raters to deal with asymmetric interactions. Ducasse and Brown's (2009) study, which collected raters' verbal reports, also reported that, since they were recording what the raters were focusing on when they were rating the co-constructed discourse in paired oral tests, their study will give valuable information concerning interactional features and descriptors which should be taken into consideration when interactional competence rating scales are developed.

Conclusion

To conclude this paper review, many further studies still need to be conducted to unravel the complexities of interactional competences and co-constructed discourse created by the students in the paired oral tests, and to create paired oral tests which are more construct valid, reliable, authentic, practical, interactive, and impactful (Bachman & Palmer, 1996), as well as to measure the interactional competences and the discourse validly and reliably.

References

Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice: Designing and developing useful language tests*. Oxford: Oxford University Press.

Bennett, R. (2012). Is linguistic ability variation in paired oral language testing problematic? *ELT Journal*, *66*(3), 337–346. https://doi.org/10.1093/elt/ccr066

Brooks, L. (2009). Interacting in pairs in a test of oral proficiency: Co-constructing a better performance. *Language Testing*, *26*(3), 341–366. https://doi.org/10.1177/0265532209104666

Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches* (4th Ed.). London: SAGE Publications.

Davis, L. (2009). The influence of interlocutor proficiency in a paired oral assessment. *Language Testing*, *26*(3), 367–396. https://doi.org/10.1177/0265532209104667

Ducasse, A. M., & Brown, A. (2009). Assessing paired orals: Raters' orientation to interaction. *Language Testing*, *26*(3), 423–443. https://doi.org/10.1177/0265532209104669

Galaczi, E. D. (2008). Peer–peer interaction in a speaking test: The case of the First Certificate in English examination. *Language Assessment Quarterly*, *5*(2), 89–119. https://doi.org/10.1080/15434300801934702

Galaczi, E. D. (2014). Interactional competence across proficiency levels: How do learners manage interaction in paired speaking tests? *Applied Linguistics*, *35*(5), 553–574. https://doi.org/10.1093/applin/amt017

Kramsch, C. (1986). From language proficiency to interactional competence. *The Modern Language Journal*, *70*(4), 366–372. https://doi.org/10.2307/326815

Laborda, J. G., Juan, N. O. de, & Bakieva, M. (2015). Co-participation in oral paired Interviews: Preliminary findings of the OPENPAU project. *Procedia - Social and Behavioral Sciences*, *191*, 559–563. https://doi.org/10.1016/j.sbspro.2015.04.614

Lazaraton, A., & Davis, L. (2008). A microanalytic perspective on discourse, proficiency, and identity in paired oral assessment. *Language Assessment Quarterly*, *5*(4), 313–335. https://doi.org/10.1080/15434300802457513

May, L. (2009). Co-constructed interaction in a paired speaking test: The rater's perspective. *Language Testing*, *26*(3), 397–421. https://doi.org/10.1177/0265532209104668

Ockey, G. J., & Li, Z. (2015). New and not so new methods for assessing oral communication. *Language Value*, *7*(1), 1–21. https://doi.org/http://dx.doi.org/10.6035/LanguageV.2015.7.1