TEST-TAKER INTERACTION – HOW DO RATERS PERCEIVE THE CONSTRUCT?

Linda Borger
University of Gothenburg

13th EALTA Conference
May, 2016, Valencia
Presentation outline

• Aim of study and context
• Background
  – Construct conceptualisation
  – Literature review
• Data and methods
• Results
• Discussion and conclusions
Aim of study and context

**Aim**
Explore the construct of peer-peer interaction from the rater’s perspective

**The test**
A paired speaking test – part of a high-stakes Swedish national EFL test at upper secondary level

**Research question**
What features of peer-peer interaction are salient to raters as they evaluate performance in a paired speaking test?
Construct conceptualisation
L2 proficiency

1. Structuralist theory (Lado, 1957)
2. Communicative competence (e.g., Canale & Swain, 1980)
3. Interactional competence (IC) (e.g., Kramsch, 1986)

• a single individual’s contribution to communication ➔
  communication as **socially situated** and **co-constructed** by
  participants (Jacoby & Ochs, 1995)

• Young (2000): interactive communication skills are ‘**local**’ and
  ‘**practice-specific**’ (highly context-dependent): “IC varies with
  the practice and with the participants” (Young, 2011, p. 428)
Construct of interactional competence: implications for rating

1. Measures of interactional competence
   **Macro**: Patterns of co-construction
     • mutuality and equality
   **Micro**: Interaction features (Hall, 1995; He & Young, 1998)
     • verbal, paralinguistic and non-verbal

2. Ratings scales
   • accounting for joint performance

3. Test-taker characteristics, task type and group size
Literature review

Conversation Analysis

interactional patterns and features (Davis, 2009; Galaczi, 2008, 2014; Lazarton & Davis, 2008; Nakatsuhera, 2011)

• collaborative, parallel and asymmetric patterns (Galaczi, 2008)
• LPI (Lazarton & Davis, 2008)

Rater report studies


• 1) interactive listening,
  2) non-verbal interpersonal communication, and
  3) interactional management

• key features of the interaction viewed as mutual achievements (May, 2009, 2011)
Data collection and analysis

Participants

31 raters
- 17 Swedish teacher-raters
- 14 external CEFR raters

Sample performances

- 12 individual audio-recorded performances from 6 paired speaking tests

Data

- Holistic scores and written comments on features that contributed to their judgement

Analyses

- Rater comments segmented and coded (VPA)
- Coding scheme developed based on the written comments and rating criteria
Results: Features attended to

- Accuracy: 22%
- Range: 19%
- Coherence: 15%
- Fluency: 13%
- Interaction: 15%
- Production strategies: 4%
- Sociolinguistic appropriateness: 1%
- Task realisation: 7%
- Intelligibility: 3%
- Other: 1%
- Produc=on strategies: 4%
- Range: 19%
- Sociolinguistic appropriateness: 1%
- Task realisation: 7%
- Intelligibility: 3%
- Other: 1%
- Produc=on strategies: 4%
- Range: 19%
- Sociolinguistic appropriateness: 1%
- Task realisation: 7%
- Intelligibility: 3%
- Other: 1%
- Produc=on strategies: 4%
- Range: 19%
- Sociolinguistic appropriateness: 1%
- Task realisation: 7%
- Intelligibility: 3%
- Other: 1%
- Produc=on strategies: 4%
Results: Evaluative response

Interaction

- 63% Positive
- 23% Negative
- 14% Mixed
Results: Raters’ orientation to interaction
Micro level

1. Topic extension moves

• He makes consistently very good contributions to the discussion, asking questions and introducing new topics. (positive) /CEFR

• Really good at adding perspectives to the male student’s topics (positive) /Sw

• He gives brief comments on his partner’s views but doesn’t use a large repertoire of strategies to keep it going on. (mixed) /CEFR
Results: Raters’ orientation to interaction
Micro level

2. Turn-taking management

• *Takes the initiative to start and to continue conversation.* *(positive) /CEFR*

• *Does not respect turns, so in a way the dialogue becomes a monologue.* *(negative) /CEFR*

• There is well timed turn taking in the conversation which gives balance in the exchange of information. Even if it’s a little like a Ping-Pong match in the beginning. *(mixed) /Sw*
Results: Raters’ orientation to interaction

Micro level

3. Interactive listening strategies

• He gets many chances to prove that he is a good listener and that he is able to vary his expressions when it comes to confirming, establishing agreements and pointing out where he agrees with the female student “same here”, “yeah, sure”, “me neither”, for instance. (positive) /Sw

• She shows she’s been listening to male partner and makes good contributions to the conversation. (positive) /CEFR
Results: Raters’ orientation to interaction  
Macro level

Asymmetric interaction
She is helpful, a bit bossy though, dominates in interaction, explains on behalf of her counterpart, overpowers rather than collaborates with him to achieve a conversation. *(mixed) /CEFR*

*(dominant)*
Actually, she controls discussion, asks the questions, asks him for clarification of what he says. Makes and helps both conversation and test flow. *(positive) /CEFR*

*(managing)*
Let’s his partner take command too often and is not as involved in the discussions as he maybe could be, which reduces the grade as it’s harder to get a full picture. *(negative) /Sw*

*(passive)*
Results: Raters’ orientation to interaction
Macro level

Co-construction of interactional patterns perceived as advantageous or disadvantageous – implications for separability of scores

• I felt that this speaker was somewhat disadvantaged due to a domineering partner. I would have liked to have heard more. /Sw

• I think she helps her partner achieve a higher grade than he has achieved before because she adapts her language and asks good questions. (Tr.) /Sw

• Hard to assess since the girl keeps interrupting and takes over ever so often. /CEFR
Results: Raters’ orientation to interaction

Mutual achievements

Comparisons

– Similarities, differences, proficiency levels, quality of interaction

1 Test-takers as individuals in relation to the rating scale

• Vocabulary not quite as comprehensive as the male speaker’s, also simpler. (negative) /CEFR

2. Test-takers’ separate contributions not clearly recognizable
Comments on quality of interaction

• The speakers help each other well here, they give and take, ask for clarifications, examples (positive) /Sw

• Interaction between the two is okay, they comment on each other but they could help each other more. (mixed) /Sw
Discussion and conclusions

• The results in line with the broad view of peer-peer interaction demonstrated in previous research (both CA and rater reports)
  – Non-verbal communication not observed

• Theoretical level: deeper understanding of the complexity of the IC construct from the rater’s perspective
  – ‘Interactionalist’ approach to construct definition (Messick, 1989; Chapelle, 1998)

• Practical level: empirical scale development (IC criterion)
  – Inform/complement the development of IC descriptors; co-constructing of performance
  – “it is clear that there was a lot more going on in the paired format than the rating scale captured” (Brooks, 2009)
  – Rater training: understand the theoretical model (cf. Orr, 2002)
Interactionalist validation of a test

In an interactionalist validation of a test, a person's performance on a test is taken to indicate an underlying trait characteristic of that person and, at the same time, the performance is also taken to indicate the influence of the context in which the performance occurs. The interactionalist definition is, in other words, a way to infer from test performance something about both a practice-specific behavior and a practice-independent, person-specific trait. Moreover, the interactionalist definition of a construct refers not only to the trait and the context but also to some theory of how the two interact.

(Young, 2011, p. 438)
Discussion and conclusions cont’d

• Further research: **Interactional patterns – relationship to scores?**
  – Two test-takers perceived as dominant awarded highest average scores in current study (cf Galaczi, 2008)

• **Co-constructed features and patterns**
  – Test-takers advantaged or disadvantaged by pairing: issue of reliability
  – In Sw system: interactional skills tested several times; enhances fairness
  – Comparisons between test-takers (cf. Orr, 2002; May, 2011) – A potentially controversial issue
Selected References


Thank you for your attention!

Linda Borger
Dept. of Education and Special Education
University of Gothenburg
linda.borger@gu.se