

Book Review:

**Investigating Pragmatics in Foreign  
Language Learning, Teaching, and Testing**

(Second Language Acquisition 30 Series)

Edited by Eva Alcón Soler and Alicia Martínez-Flor (2008)

Bristol, Buffalo, and Toronto: Multilingual Matters

ISBN: 978-1-84769-085-5 (pbk.) 978-1-84769-085-2 (hbk.)

Pbk: ¥5,560 JPY, £27.95 GBP, \$54.95 USD



Weaving together information about the learning, teaching, and testing of pragmatics in foreign language contexts, this book provides varied perspectives on developmental pragmatics. Essentially it raises two questions: (1) how should pragmatics be taught? and (2) how should it be tested? This review focuses only on the latter issue, to which nearly a third of the book's 12 chapters are devoted. Four broad assessment methods and four analytical orientations from this 266-page text are outlined in this review.

***Pragmatic Assessment Methods***

***Method 1: Discourse Completion Tasks***

The most common way to get pragmatic data is through discourse completion tasks (DCTs). All DCTs consist of a prompt, response, and sometimes a rejoinder (which might be considered a back-shifted prompt).

Typical prompt formats are written passages, video clips, or pictures. Yamashita (p. 215) regards the first option as inferior to latter ones due to a tendency of many foreign language examinees to misunderstand parts of text-based prompts and/or use snippets from such prompts to construct their responses.

DCT response formats consist of either of constructed responses (e.g. "Mention what you would probably say now") or fixed responses (e.g. "Select the most appropriate choice, A-D").

DCT response modalities can be either oral or written. It is important to recognize the differences between these modalities and avoid eliciting responses for one modality via the other (e.g. "Write down what you would imagine saying"). An otherwise outstanding study by Eslami and Eslami-Rasekh in this volume had this weakness.

Despite criticisms of DCTs by numerous researchers, they appear to be a valuable research tool, especially when used in conjunction with other instruments. Brown's review of several well-known pragmatic studies employing DCTs in Chapter 11 revealed that some were problematic (with KR-21 ranges of .45 - .62), but most were moderately reliable (with .76 - .92 KR-21 ranges).

#### *Method 2: Role-plays*

Role-plays have the advantage of testing pragmatic situations in full discourse contexts, but they are generally time-consuming to administer. Though costly to produce, Yamashita offers high praise for video-mediated role-plays (p. 214). Indeed, when comparing data from Ahn's (2005) study about the effectiveness of videotaped role-plays vis-à-vis oral and written DCTs and video-mediated self-reflections, Brown demonstrated that the former instrument was the most dependable for the data provided according to G-theory (p. 243).

#### *Method 3: Self-assessments and think aloud protocols*

How well can FL students rate their own performance? Data from Brown's analysis of Ahn's (2005) study of Korean language learners indicates adult foreign language learners were able to rate videotaped clips of their own pragmatic performances according to a 5-point Likert scale in a way that correlated not so closely with the 3 native-speaker raters (Cohen's  $d = .95$ ). This indicates a significant magnitude of difference between the two groups: there was less than 45% of overlap between the two ratings. When considering self-ratings we must also reflect on what incentive, if any, would there be for examinees to inflate or deflate their scores. Moreover, when using self-ratings, it is best to incorporate other data sources into the research design for corroboration.

To gain possible insights into the mental processes foreign language learners undergo while engaging in pragmatic tasks, verbal reports may also be an option. In this text, the verbal reports of two groups of Australian learners of Indonesian are considered in depth. If we accept the underlying assumption that retrospective verbal reports provide accurate insights into what's occurring in learners' minds, then it appears that elementary level language learners devote more time to verbal planning (i.e. deciding how to structure a conversation) than proficient learners do. Since all of the proficient foreign language students in the study mentioned in this book spent a year or more in Indonesia, the author speculates

living abroad was the reason for their enhanced pragmatic skills (p. 90). Unfortunately, the research design didn't permit such a conjecture. To ascertain *why* inter-group pragmatic ability differs, a longitudinal design with controlled and experimental variables would be required.

#### *Method 4: Translation tasks*

An argument for reintroducing translation into foreign language classes is offered by Juliane House. Acknowledging the limitations of the traditional grammar-translation method and criticizing the linguistic manipulation of isolated sentences, House contends that "communicative translation" practices are not only possible, but in fact a useful way to enhance pragmatic awareness. She suggests translation should be taught as a "5<sup>th</sup> language skill" at the upper levels. In spite of the hegemony of monolingual theory in EFL instruction, House shrewdly notes how translation exercises often appear in university entrance exams because they are, in her opinion, efficient ways to evaluate many comprehension skills. Since her claims have no supporting data, they were less than convincing. Nonetheless, House seems on the mark in acknowledging that a keen understanding of pragmatic and cross-cultural issues is essential for effective translation.

#### *Analytical Orientations*

Three of the studies from this text are grounded in classical test theory and Brown provides a coherent explanation of how the classical notion of test reliability can be influenced by sample size according to the Spearman-Brown Prophecy Formula. Four of the studies here employ qualitative techniques, and it is interesting to see how well (or poorly) the data for those studies is triangulated. A non-technical introduction to G-theory is also offered and Carsten Roever's final essay reveals some of the strengths of the Multi-facet Rasch Measurement Model (MFRM).

What might confuse some readers is which research orientation described in this book is most appropriate for their needs. When should qualitative orientations be favored over quantitative ones? How well can qualitative and quantitative data be combined? Also, is classical test theory now entirely passé, as some Rasch adherents claim (Halkitis, 1993, par. 1; Prieto, Alonso, and Lamarca, 2003)? Finally, what are the respective advantages of G-theory and MFRM? The editors of this work express no preferences. Many readers in the testing field, however, may want guidance concerning such questions.

### **Conclusion**

One of the best things about this book is how it raises incisive questions about ongoing pragmatic research. For example, Yamashita echoes Roever's 2001 concern as to whether adequate pragmatic component samples can be practically obtained in a test (p. 217). Yamashita also astutely emphasizes how tests of pragmatic proficiency need to include both productive and comprehension tasks, since many foreign language learners display a gap between what they can understand and what they can produce. González-Lloret also remarks how pragmatic performance tends to vary over time and how long-term longitudinal data is needed to obtain accurate accounts of performance (p. 121).

This book is an attempt to wed interlanguage pragmatic studies with four different analytical orientations. It is a curious hybrid, but the fact that it lacks any thematic index and many of its chapters do not fit into its overall structure makes it difficult to navigate through. Despite these shortcomings, parts of this text are informative. Those wishing to get a better grasp of recent pragmatic testing studies will find the concluding four chapters of particular value.

- reviewed by Tim Newfields

Toyo University Faculty of Economics

### **Works Cited**

- Ahn, R. C. (2005). Five measures of interlanguage pragmatics in KFL (Korean as a foreign language) learners. Unpublished Ph.D. thesis, University of Hawaii at Manoa.
- Halkitis, P. N. (1993). Computer-adaptive testing algorithm. *Rasch Measurement Transactions*, 6 (4). Retrieved March 30, 2009 from [www.rasch.org/rmt/rmt64cat.htm](http://www.rasch.org/rmt/rmt64cat.htm)
- Prieto, L. Alonso, J., Lamarca, R. (2003, July 28). Classical test theory versus Rasch analysis for quality of life questionnaire reduction. *Health Quality Life Outcomes*, 1 (27). Retrieved March 30, 2009 from <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=194220>
- Roever, C. (2001). A web-based test of interlanguage pragmalinguistic knowledge: Speech acts, routines, and implicatures. Unpublished Ph.D. thesis, University of Hawaii at Manoa.