

## Contents

### Articles

- Towards a computer-delivered test of productive grammatical ability 443  
*Carol A. Chapelle, Yoo-Ree Chung, Volker Hegelheimer, Nick Pendar and Jing Xu*
- The effect of the multiple-choice item format on the measurement of knowledge of language structure 471  
*Michael Currie and Thanyapa Chiramanee*
- The effect of the use of video texts on ESL listening test-taker performance 493  
*Elvis Wagner*
- Explaining ESL essay holistic scores: A multilevel modeling approach 515  
*Khaled Barkaoui*
- The challenge of validation: Assessing the performance of a test of productive vocabulary 537  
*Tess Fitzpatrick and Jon Clenton*
- The place of language testing and assessment in the professional preparation of foreign language teachers in China 555  
*Yan Jin*
- Interaction in group oral assessment: A case study of higher- and lower-scoring students 585  
*Zhengdong Gan*
- Sensitivity of narrative organization measures using narrative retells produced by young school-age children 603  
*John Heilmann, Jon F. Miller and Ann Nockerts*

## Book review

- Understanding language testing 627  
*D. Douglas, reviewed by Fred Davidson*

## Test review

- Canadian Academic English Language (CAEL) Assessment 631  
*Reviewed by Margaret E. Malone*

- List of reviewers 637  
Errata 639  
Corrigendum 640

[www.languagetestingjournal.com](http://www.languagetestingjournal.com)

## Towards a computer-delivered test of grammatical ability

Carol A. Chapelle, Yoon Hegeheimer, Nick Pe  
Iowa State University, USA

### Abstract

This study piloted test items that measure productive grammatical ability in foreign language learners' development of morphological complexity to create a framework of grammatical ability. Results from four pilot test administrations are reported, including difficulties and developmental stages of test scores in relation to productive ESL grammatical ability and of moving forward with computer-delivered tests.

### Keywords

constructed responses, interpretation, language acquisition, validity argument

[L]anguage testers can be criticized for using discrete-point tasks of grammatical ability to measure grammatical development that does not reflect a model of grammatical ability; rather, it reflects a model of 'communicative effectiveness'.

Purpura's statement is relevant to the design of foreign language (ESL) student tests. The model equated with out-of-favor multiple-choice tests, has been eliminated, and replaced by a writing prompt. Examinees are evaluated on a rubric consisting of descriptors of grammatical ability across courses. The evaluations consist of 'not, and if not, which of the ESL'.

### Corresponding author:

Carol A. Chapelle, TESL/Applied Linguistics  
Email: [carolc@iastate.edu](mailto:carolc@iastate.edu)