

feedback has an effect beyond the immediate composition, thus helping to improve long-term writing ability.

**00-509 Porte, Graeme** (Universidad de Granada, Spain; *Email*: gporte@platon.ugr.es). Where to draw the red line: error toleration of native and non-native EFL faculty. *Foreign Language Annals* (New York, USA), **32**, 4 (1999), 426–34.

The central purpose of the study reported here was to investigate professors' reactions to the academic writing of non-native students. The object was to ascertain any statistically significant differences between the error-gravity perceptions of native-speaker and non-native speaker faculty, and what such findings might reveal about the respective error tolerance of these professors. Fourteen native and sixteen non-native university professors were asked to respond to student errors based on a random selection of 54 anonymous ungraded compositions written by students at Granada University. General comparison of the error grading between the two groups of participants revealed small differences in the error toleration of native and non-native faculty. While differences do exist in the perceived gravity of specific errors, it would appear that teachers in this study generally agreed in their judgments. Nevertheless, there was evidence that errors are not being perceived as seriously as one would have expected and the implications of this finding are discussed.

**00-510 Ruiz-Funes, Marcela** (East Carolina U., Greenville, NC, USA). Writing, reading, and reading-to-write in a foreign language: a critical review. *Foreign Language Annals* (New York, USA), **32**, 4 (1999), 514–26.

The development of reading and writing skills in a foreign language (FL) has a central role in the FL curriculum. The importance of these abilities is even stronger in upper-level courses where reading and writing are assigned in combination, as in the case of reading-to-write tasks. In these tasks students are asked to read articles or literary selections and to react and respond to them in an insightful and critical manner. The complexity involved in these skills as well as the web of processes that readers/writers orchestrate have been unveiled by researchers in both first and second/foreign languages seeking an understanding of the cognition of reading-to-write acts. In this article a review is presented of the most influential work carried out on the process of reading-to-write in the last two decades. Reference is also made to the impact that research on writing process and reading process has had on the FL profession.

**00-511 Victori, M.** (Universitat Autònoma de Barcelona, Spain). An analysis of writing knowledge in EFL composing: a case study of two effective and two less effective writers. *System* (Oxford, UK), **27**, 4 (1999), 537–55.

The study reported here seeks to throw some light on a relatively untouched area, by analysing how differences

in the beliefs or metacognitive knowledge (MK) held about writing relate to differences in English as a Foreign Language (EFL) writing skills. Data were collected from four undergraduate Spanish students—two good and two poor writers—enrolled in EFL classes at the University of Barcelona. They were first required to take an English test and write an argumentative essay to assess their language and writing proficiency; and were subsequently interviewed and required to think aloud as they wrote another such essay. The study revealed a number of areas where the knowledge of the two pairs clearly differed. On the whole, these differences pointed to a more appropriate and comprehensive view of the writing process, which they were able to apply more flexibly. In contrast, the less successful writers' MK was limited and inadequate. The case studies also revealed the clear relationship that exists between the writers' MK and the strategies they deployed, underscoring the major role played by metacognitive knowledge in providing a rationale for the learners' approach to writing, as well as giving researchers a more thorough understanding of the learners' writing process.

## Language testing

**00-512 Bachman, Lyle F.** (U. of California, LA, USA; *Email*: bachman@humnet.ucla.edu). Modern language testing at the turn of the century: assuring that what we count counts. *Language Testing* (London, UK), **17**, 1 (2000), 1–42.

In the past twenty years, language testing research and practice have witnessed the refinement of a rich variety of approaches and tools for research and development, along with a broadening of philosophical perspectives and the kinds of research questions that are being investigated. While this research has deepened our understanding of the factors and processes that affect performance on language tests, as well as of the consequences and ethics of test use, it has also revealed lacunae in our knowledge, and pointed to new areas for research. This article reviews developments in language testing research and practice over the past twenty years, and suggests some future directions in the areas of professionalising the field and validation research. It is argued that concerns for ethical conduct must be grounded in valid test use, so that professionalisation and validation research are inseparable. Thus, the way forward lies in a strong programme of validation which includes considerations of ethical test use, both as a paradigm for research and as a practical procedure for quality control in the design, development and use of language tests.

**00-513 Gruba, Paul** (U. of Melbourne, Australia). Key barriers in the development of computer-based testing. *Prospect* (Macquarie U., Sydney, Australia), **14**, 3 (1999), 46–53.

Despite the widespread interest in computer-assisted language learning (CALL), the use of computers for the purposes of language assessment has lagged behind advances made in teaching. This article argues that the lack of development in computer-based testing is the result of four factors: (a) a lack of interest amongst CALL researchers and testing professionals; (b) persistent technical limitations of computers; (c) unresolved theoretical concerns; and (d) equity issues. Each of these factors is discussed in turn, and it is concluded that efforts to improve and evaluate CALL programs will be hindered by the lack of development of computer-based assessment instruments. The article concludes with suggestions for how computer-based testing research and development may be fostered.

**00-514 Jansen, Heleen** (Kaj Munkcollege, Hoofddorp, The Netherlands) and **Peer, Caroline**. Centrale Examens tekstbegrip moderne vreemde talen havo vwo en het gebruik van woordenboeken. [Using dictionaries with national foreign-language examinations for reading comprehension.] *Levende Talen* (Amsterdam, The Netherlands), **544** (1999), 639-41.

The Dutch examination authorities have approved the use of dictionaries with foreign-language examinations for reading comprehension. This decision was partly based on a small-scale field study and a literature review. The authors here report on a larger study among 200 secondary school pupils recruited from 18 different schools. The pupils were allowed to use a dictionary with a French examination, three to four months prior to the actual exams. The pupils were randomly assigned to two groups: one group was allowed to use a dictionary, the other was not. The results show that using a dictionary does not influence test scores in a significant way. By contrast, pupils are very positive about the use of dictionaries: it strongly reduces their test anxiety, and they claim they better understand the texts when using a dictionary.

**00-515 Jafarpur, Abdoljavad** (Shiraz U., Iran). What's magical about the rule-of two for constructing C-tests? *RELC Journal* (Singapore), **30**, 2 (1999), 86-100.

C-tests are often claimed to be the best in the family of tests of reduced redundancy. They are generally regarded as theoretically and empirically valid and reliable measures of language ability. A C-test contains four to six texts and a total of 100 items. It is constructed according to the rule-of two, which involves deleting the second half of every other word beginning from the second word of the second sentence. The study reported here investigates five versions of a C-test and a standard cloze test with 340 Iranians majoring in English. The C-tests were constructed with three different deletion starts and two different ratios. The results suggest that there is nothing magical about the rule-of two; other deletion rates and deletion starts yielded more or less similar results. The paper concludes with suggestions for improving the C-test.

**00-516 Lewkowicz, Jo A.** (U. of Hong Kong; *Email: jolewkow@hkusua.hku.hk*). Authenticity in language testing: some outstanding questions. *Language Testing* (London, UK), **17**, 1 (2000), 43-64.

This article is divided into two main sections. Following the introduction, Section II takes a look at the concept of authenticity and the way this notion has evolved in language testing and more recently in general education. It argues that, although our understanding of the notion of authenticity has developed considerably since it was first introduced into the language testing literature in the 1970s, many questions remain unanswered. In an attempt to address one of the outstanding issues, Section III presents a study looking at the importance of authenticity for test takers. It shows that test takers are willing and able to identify the attributes of a test likely to affect their performance. However, these attributes do not necessarily include authenticity, which has hitherto been considered an important test attribute for all stakeholders in the testing process. The article concludes that much more research is needed if the nature and role of authenticity in language testing are to be fully understood.

**00-517 Sasaki, Miyuki** (Nagoya Gakuin U., Japan; *Email: f43689g@nucc.cc.nagoya-u.ac.jp*). Effects of cultural schemata on students' test-taking processes for cloze tests: a multiple data source approach. *Language Testing* (London, UK), **17**, 1 (2000), 85-114.

The study reported here investigated how schemata activated by culturally familiar words might have influenced students' cloze test-taking processes. Sixty Japanese English as a Foreign Language students were divided into two groups with equivalent English reading proficiency. They completed either a culturally familiar or an unfamiliar version of a cloze test. Partially replicating Chihara *et al.*'s (1989) experiment, several unfamiliar words in the original cloze test passage were changed to more familiar ones in the modified version. Unlike in the original experiment, however, students were asked to give verbal reports of their test-taking processes, and to recall the passage after they had completed the tests. Thus, the two groups' test-taking activities were compared in terms of: (1) item performance; (2) expressing correct understanding of the key terms while solving the items and recalling; (3) the amount of text information they used to complete the items; and (4) the quantities and qualities of the final recalls. Results demonstrated that those who read the culturally familiar cloze text tried to solve more items and generally understood the text better, which resulted in better performances than those of the students who read the original text. These results also support the claim that cloze tests can measure higher-order processing abilities.

**00-518 Vermeer, Anne** (Tilburg U., The Netherlands; *Email: anne.vermeer@kub.nl*). Coming to grips with lexical richness in spontaneous

speech data. *Language Testing* (London, UK), **17**, 1 (2000), 65–83.

In spontaneous speech data, lexical richness is generally operationalised by measures in which the relation between the number of types and tokens plays a role, of which the Type/Token Ratio (TTR) is the most famous. This article discusses the reliability and validity of different measures of lexical richness in various language data research and computer simulations, and examines the behaviour of these measures in spontaneous speech data of first language and second language children learning Dutch, aged four to seven, compared with their lexical abilities as measured by tests. The results show that neither the validity nor the reliability of the measures were satisfactory, especially the widely applied TTR. Initially, the number of types, or lemmas, and the Guiraud and Uber indexes seem to be adequate measures. However, in later stages of vocabulary acquisition (from 3000 words on), neither is valid. It is suggested that more effective measures of lexical richness might be based not on the distribution of or the relation between the types and tokens, but on the degree of difficulty of the words used, as measured by their (levels of) frequency in daily language input.

## Teacher education

**00-519 Amores, María J.** (West Virginia U., USA). Preparing graduate teaching assistants: an investment in excellence. *Foreign Language Annals* (New York, USA), **32**, 4 (1999), 441–68.

In the beginning and intermediate language programmes at most major universities, graduate teaching assistants (GTAs) are a major factor in determining the overall effectiveness of the instruction provided. Based upon the belief that excellence in teaching is crucial to the development of functional ability in a foreign language, the primary goal of GTA Coordinators must be to assure the quality of the instruction in the undergraduate courses for which the Coordinator is held accountable. This article discusses the elements of one programme established to help GTAs make the transition from language student to language teacher, and to assure that a high quality of instruction is maintained.

**00-520 Farrell, Thomas S. C.** (Nat. Inst. of Ed., Singapore). The reflective assignment: unlocking pre-service English teachers' beliefs on grammar teaching. *RELC Journal* (Singapore), **30**, 2 (1999), 1–17.

Recently, there has been a call for teacher education programmes to acknowledge the influential role of student teachers' prior knowledge and personal understandings. This is especially important given that, in many countries around the world, the emphasis in English teaching methodology has shifted over the

years from overt grammar instruction to a more communicative approach. Many pre-service teachers have been educated in English by traditional methods of drill and memorisation of grammar rules, so there can be a mismatch between their own past learning experiences and their current teacher education programme. An important question then arises as to how these prior experiences, often tacit, can be made more conscious and integrated into the curriculum. This paper sets out to show how pre-service teachers' beliefs in Singapore were unlocked by the use of a three-part reflective assignment. The paper begins with a brief discussion of teachers' beliefs. The study is next outlined: the course, the assignment and the student teachers' past experiences are described, and a detailed analysis of responses from five of the 34 participating teachers is presented. Finally, the same student teachers' reflections of their actual teaching experiences are outlined and discussed.

**00-521 Murdoch, George** (United Arab Emirates U., Al Ain; *Email*: time@emirates.net.ae). Introducing a teacher-supportive evaluation system. *ELT Journal* (Oxford, UK), **54**, 1 (2000), 54–64.

This article begins by examining the key principles underpinning a supportive approach to teacher evaluation. These principles highlight the need to ensure that evaluation plays a part in establishing an institutional concern for teacher development and teacher support. A number of elements are recommended for creating such a system. These are described in terms of their functioning in the teacher-performance review system which has been developed for the General Requirements Unit English Program at the author's institution. In the final part of the paper, the results of a survey of teachers' views on the system are presented and discussed. These suggest that teachers generally find the system supportive and effective. The description of this system is intended to interest teachers, managers and administrators in other institutions looking to introduce more motivating and more progressive performance review procedures.

**00-522 Tanner, Rosie, Longayroux, Désirée, Beijaard, Douwe and Verloop, Nico** (U. of Leiden, The Netherlands; *Email*: rosie.tanner@wxsl.nl). Piloting portfolios: using portfolios in pre-service teacher education. *ELT Journal* (Oxford, UK), **54**, 1 (2000), 20–30.

The promotion of reflection is nowadays considered to be an important goal in teacher education, because it is assumed that reflection plays a major part in the development of teachers. The present authors support the idea that portfolios can help to develop this vital skill of reflective thinking. This article recounts their experiences with using portfolios as an instrument for professional development during a one-year pre-service teacher education course for language graduates. After sketching the context of their work, they describe why and how they are working with portfolios. They illustrate their belief that portfolios can (a) demonstrate a trainee's learning process over time, (b) illustrate an individual's