

ciency on entry were established through matriculation examination scores; and their English achievement through their proficiency test scores on completion of the two-year intensive English programme. 16 learner variables were established, and a hypothetical causal model was constructed. Relationships within the model and between learner variables and achievement were examined through Partial Least Squares analysis, and it was found that six variables had direct effects. Three were traits that existed prior to students' admission, i.e. Sex, and Proficiency in first and second language. The other

three were formed by clusters of strategies relating to Vocabulary learning, Tolerating ambiguity (Risk-taking), the only negative direct effect, and Mother-tongue avoidance. Management strategies had the strongest direct effect on English achievement. Finally, the direct effects of belief variables on strategy variables were examined and found to be strong and consistent. Selective qualitative data are used throughout to provide further insight into and illustrations of differences in strategy use between more and less successful learners.

97-377 Zhao, Yong (Michigan State U.). The effects of listeners' control of speech rate on second language comprehension. *Applied Linguistics* (Oxford), **18**, 1 (1997), 49-68.

Speech rate has been identified as a major factor affecting listening comprehension. Despite the common-sense belief that slower rates facilitate listening comprehension, empirical studies have yielded contradictory findings. A review of previous studies of speech rate and listening comprehension revealed methodological problems. Using a novel approach and recent developments in computer technology, this study examined the issue of speech rate and lis-

tening comprehension from a different perspective. Subjects were 15 non-native speakers of English at a U.S. university. By giving the control of speech rate to the students and by attending to individuals instead of groups, this study concluded that (a) when given control, students' listening comprehension improved and (b) improved listening comprehension was achieved by slowing down the speech rate.

Language testing

97-378 Chalhoub-Deville, Micheline (U. of Minnesota). Theoretical models, assessment frameworks and test construction. *Language Testing* (London), **14**, 1 (1997), 3-22.

This article reviews the usefulness of several models of proficiency that have influenced second language testing in the last two decades. The review indicates that several factors contribute to the lack of congruence between models and test construction, and makes a case for distinguishing between theoretical models, which attempt to represent the proficiency

construct in various contexts, and operational assessment frameworks, which depict the construct in particular contexts. Additionally, the article underscores the significance of an empirical, contextualised and structured approach to the development of assessment frameworks.

97-379 Cheng, Liying (U. of Hong Kong). How does washback influence teaching? Implications for Hong Kong. *Language and Education* (Clevedon), **11**, 1 (1997), 38-54.

There is some evidence to suggest that tests have washback effects on teaching and learning. The extensive use of test scores for various educational and social purposes in society nowadays has made the effect of washback a distinct educational phenomenon. This paper presents preliminary research findings on the washback effect of the Hong Kong Certificate of Education Examination (HKCEE) in English in Hong Kong secondary schools by employing various methodological techniques such as two questionnaires, interviews and classroom observations in sampled schools. It further discusses

the nature of washback effect, the major teaching and learning aspects influenced by it, the different stages of washback effect, and the types of washback effects observed. Preliminary results indicate that washback effect works quickly and efficiently in bringing about changes in teaching materials, which is due largely to the highly adaptable and commercial nature of Hong Kong society, and slowly and reluctantly and with difficulties in the methods teachers employ. It is suggested that the latter effect may be caused by the constraints imposed upon teaching and teachers in the present schools.

97-380 Giménez, Julio César (Nat. U. of Córdoba). Process assessment in ESP: input, throughput and output. *English for Specific Purposes* (Oxford), **15**, 3 (1996), 233–41.

In most English for Specific Purposes (ESP) programmes students' performance tends to be evaluated following the well-established practice for testing only end-products, ignoring any assessment of the learning process. The author sees the most serious drawback to this as the end-product evaluation coming too late in the learning process for formative feedback to occur. This paper discusses the

importance of implementing *process assessment* in ESP courses, and suggests that, since this alternative procedure evaluates not only students' 'products' but also the way in which they have come to acquire the proficiency needed to produce them, it offers the opportunity to improve outcomes while there is still time to do so. The paper gives examples drawn from a local programme at tertiary level.

97-381 Grant, Leslie (Central Michigan U.). Testing the language proficiency of bilingual teachers: Arizona's Spanish proficiency test. *Language Testing* (London), **14**, 1 (1977), 23–46.

Currently in the USA, a total of 28 states offer bilingual teacher certification and/or endorsement. Of these states, 17 require that the prospective teachers demonstrate proficiency in their second language, in English, or in both languages. In the context of nationwide bilingual teacher testing procedures, this article describes a criterion-referenced test for bilingual endorsement used in Arizona. Test scores and

questionnaires provided information with respect to classification dependability and rater consistency. Examinee questionnaire responses, teacher questionnaire responses and test section analysis all contributed evidence for validity. As a result of this study, further research on testing teachers' language proficiency is recommended and considerations for bilingual teacher training are discussed.

97-382 Klein-Braley, Christine (U. of Duisburg). C-Tests in the context of reduced redundancy testing: an appraisal. *Language Testing* (London), **14**, 1 (1997), 47–84.

In language testing, the concept of reduced redundancy has been a fruitful approach for the development of major test procedures. The way in which examinees perform under conditions of 'noise' is believed to provide evidence for the level of their current status in overall or general language proficiency. This article reports an investigation comparing the empirical performance of C-Tests with other representatives of the 'family' of reduced redun-

dancy tests: classical cloze, cloze-elide, multiple-choice cloze. The criterion for empirical validity is DELTA, the Duisburg English Language Test for Advanced Students. Overall, the C-Test emerges as the most economical and reliable procedure; it has the highest empirical validity and is shown to be the best representative of the general factor in the battery.

97-383 Richards, Brian (Reading U.) and **Chambers, Francine** (Southampton U.). Reliability and validity in the GCSE oral examination. *Language Learning Journal* (Rugby), **14** (1996), 28–34.

This article examines the reliability and validity of teacher assessments in the particularly problematic area of open-ended oral production tasks, focusing on the UK GCSE (General Certificate in Secondary Education) Higher Level conversation in French. The conversations of 30 GCSE candidates were marked on two occasions by 24 secondary teachers, and on one occasion by 13 trainee teachers. Three assessment procedures were used: Scheme A, criterion-related global impression; B, categorical with descriptors; and C, categorical without descriptors. In terms of reliability, results indicated huge differences in teacher consistency; there was no correlation between self-consistency and years' of

experience, however, with quality rather than quantity of experience proving important. Ability to mark consistently was significantly affected by the mark scheme; Schemes A and C proved superior in this respect to Scheme B. Two areas were investigated in terms of validity: range of vocabulary and complexity of structure, the latter proving much more difficult for the teachers. It is suggested that accurate assessment of oral performance remains hampered by the apparent arbitrariness of current evaluative criteria, which are neither derived from a principled model of language proficiency nor from an analysis of communicative development.

97-384 Wigglesworth, Gillian (Macquarie U.). An investigation of planning time and proficiency level on oral test discourse. *Language Teaching* (London), **14**, 1 (1997), 85-106.

The inclusion of planning time in semi-direct oral interaction tests adds considerably to the overall length of the test, and it is important to be clear that the increase in length is justified by the language outcome. Previous research has shown that the effect of planning time in second language can differentially influence the resultant discourse, with planned discourse eliciting more complex language on a range of measures. However, where planning time has been provided it has generally been a substantial amount of time (ten minutes or more), and in a second language classroom situation, rather than a testing situation. Where planning time is provided in an oral interaction test it is generally limited to

one or two minutes. In this study planning time was manipulated as a variable in a trial administration of a semi-direct test. Discourse analytic techniques were then used to determine the nature and/or significance of differences in the elicited discourse across the two conditions in terms of complexity and accuracy. In addition, candidates were divided into high-proficiency versus low-proficiency groups. The findings suggest that, for the high-proficiency candidates, planning time may improve accuracy on some measures where the cognitive load of the task is high, but that this effect does not extend to the low-proficiency candidates.

Teacher education

97-385 LeBlanc, Leona B. (Florida State U.). Testing French teacher certification candidates for speaking ability: an alternative to the OPI. *The French Review* (Champaign, IL), **70**, 3 (1997), 383-94.

Many state departments of education require that the French teachers demonstrate their skills and competencies by passing a test in order to be certified to teach in K(indergarten) through to 12. One such test is the Florida Initial Teacher Subject Area Test (ITSAT), which includes an assessment of

speaking ability. A discussion of this test, which was developed as an alternative to the Oral Proficiency Interview (OPI), is presented. In addition, the paper includes the results of a reliability study on this speaking test and suggests various applications of this assessment tool.

97-386 Pacek, Dorota (U. of Birmingham). Lessons to be learnt from negative evaluation. *ELT Journal* (Oxford), **50**, 4 (1996), 335-42.

In-service training (INSET) programmes which introduce novel approaches to teaching often seem exciting and eye-opening while the course lasts. However, the programmes do not always take account of what will happen when the course participants return to their daily routines. This article describes the post-course evaluation of an INSET programme for Japanese secondary school teachers of English, which found that the communicative approach introduced in the course was not implemented by a number of the participants. This reluctance

to innovate seems to stem from differences between British and Japanese educational and cultural traditions. It is suggested that INSET courses should therefore present a range of traditional and modern approaches to English as a foreign language methodology, so that teacher trainees can select those which are most appropriate in their particular circumstances. Such courses should be more sensitive to the participants' cultural and educational backgrounds, and provide a cognitive element from the start.

97-387 Robinson, John and Heyes, Irene (Manchester Metropolitan U.). Conflicting models of teacher training in multi-ethnic classrooms: journal of a mentor. *Language, Culture and Curriculum* (Clevedon), **9**, 2 (1996), 120-32.

The paper examines part of a journal kept by a mentor (i.e. an experienced teacher supervising the teaching practice of a student teacher) in a multi-

ethnic classroom in Britain. The mentor was an expert in the bilingual education of ethnic minority children, and the journal describes difficulties she