development of a dissertation writing support program for ESL graduate research students. *English for Specific Purposes* (Exeter, UK), **17**, 2 (1998), 199–217.

Despite an explosion in the number of students writing graduate theses in a language other than their first, there are very few accounts either of research into the difficulties encountered by these students or of writing programmes designed to help such students present dissertations written to an acceptable standard. This article describes and evaluates a programme developed within the English Centre at the University of Hong Kong to assist students who are required to present dissertations in English. The programme was based on data collected from detailed interviews with graduate supervisors and a survey of graduate students, as well as an analysis of extended pieces of graduate writing.

**99–134 Bilash, Olenka S. E.** (U. of Alberta, Canada). Planning for writing instruction in a middle-years immersion/partial immersion setting. *Foreign Language Annals* (New York), **31**, 2 (1998), 159–68.

Based on a literature review and the reflective practice of a second language teacher educator, this paper discusses three categories of planning for writing instruction for middle-years students in an immersion or partial immersion context: planning for long-term balance, planning for daily implementation, and planning for task success. It highlights five considerations that may help reduce student anxiety and improve both the experience of writing and the product: selecting topics, making the writing task as authentic and communicative as possible, providing the language supports necessary for success in a second language context, choosing forms with an appropriate amount of text, and choosing forms governed by an appropriate number of rules. A template for planning writing tasks is included, along with several examples.

**99–135 Lee, Icy** (Douglas Coll., BC, Canada). Enhancing ESL students' awareness of coherence-creating mechanisms in writing. *TESL Canada Journal / La Revue TESL du Canada* (Burnaby, B.C.), **15**, 2 (1998), 36–49.

Coherence has long been regarded as central to effective writing. However, despite the recent advances in textual linguistics which have contributed to a broader understanding of coherence, the concept is still not well understood; and it is often regarded as a fuzzy and elusive concept that is difficult to teach and difficult to learn. This article attempts to identify the text-based and reader-based features which are crucial to the creation of coherence, and describes a number of classroom tasks which can be exploited to enhance English as a Second Language (ESL) students' awareness of coherence-creating mechanisms in writing.

**99–136 Shaw, Philip** (Arhus School of Business) **and Liu, Eric Ting-Kun**. What develops in the development of second-language writing? *Applied Linguistics* (Oxford, UK), **19**, 2 (1998), 225–54.

Counts of register features have produced useful data on varieties of written English and on differences between writers of differing writing and language proficiency. This article reports an attempt to extend this procedure to an evaluation of the same foreign-language writers at two different stages. Participants were incoming overseas university students taking a full-time summer course in English for Academic Purposes–with an emphasis on writing–prior to beginning post- or undergraduate studies. The study shows that the procedure can differentiate the products of developing writers at two relatively close points in time, and that a more detailed examination of significant changes can be revealing about patterns of learning. The major changes were from features of spoken English to those more typical of formal writing, both in surface detail and in more fundamental characteristics. There was less change in complexity of construction or variety of vocabulary. Improved correctness in the structures used was balanced by errors in new structures being attempted. The participants had been discriminating in their acceptance of academic style and actively sensitive to genre and other requirements.

**99–137 Sullivan, Kirk** (Umea U., Sweden), **Kollberg, Py and Palson, Eva**. Trace-It. *Babel (AFMLTA)* (Victoria, Australia), **33**, 1 (1998), 22–27.

This paper describes *Trace-It*, a computer-based tool that traces the writing process and which is in the early stages of adaptation from linguistic research to the modern language teaching and learning environment. The program is designed as a user-friendly tool based on keystoke logging. Its goal is to make accessible extra information about an individual student's written second language skills. Once identified, these features can be used in the planning of a particular learning program. A number of advantages of the program are identified, and examples demonstrate how its use facilitates the study of the revision process by following a writer's continual reshaping of a text. The second part of the paper describes a small-scale pilot study which used the program to examine differences in revision behaviour as a group of Swedish students' ability in written English improved. Findings showed that *Trace-It* provides useful additional information about the writing process that can be used by both students and teachers to identify strengths and weaknesses in written proficiency. It is suggested that, in its present form, *Trace-It* is more useful on a one-to-one tutoring basis than as an effective classroom tool on a regular basis.

# Language testing

**99–138 Buck, Gary and Tatsuoka, Kikumi** (Educational Testing Service, Princeton, NJ, USA). Application of the rule-space procedure to language



https://doi.org/10.1017/S0261444800013628 Published online by Cambridge University Press

testing: examining attributes of a free response listening test. *Language Testing* (London), **15**, 2 (1998), 119–57.

The study reported here uses the rule-space methodology to explore the cognitive and linguistic attributes that underlie performance on an open-ended, short-answer, listening comprehension test. This is apparently the first application of the methodology to language testing. The rule-space methodology is an adaptation of statistical pattern recognition techniques applied to the problem of diagnosing the cognitive attributes (knowledge, skills, abilities, strategies, etc.) underlying test performance. The methodology provides diagnostic information about the individual test-takers on each of these attributes. Based on a literature search, attribute candidates were identified which seemed likely to explain performance on the listening test. Two rule-space analyses were carried out, and the final attribute list had 15 attributes, with 14 interactions. With this, 96% of the test-takers were successfully classified into their respective latent knowledge states, and given scores on each attribute. This result suggests that the rule-space methodology can be used to identify attributes underlying performance on language tests. The study also provided useful information about the listening construct.

**99–139 Conlam, David** (The Chinese U. of Hong Kong). Interactive evaluation of listening comprehension: how the context may help. *Computer Assisted Language Learning* (Lisse, The Netherlands), **11**, 1 (1998), 35–53.

This paper describes the design and implementation of a computer-based listening test–the 'Text Dictation'. The paper first discusses the importance of listening centring on a coherent text, rather than short listening fragments, and the drawbacks of listening programs where the required responses consist of non-interactive test types such as appropriate-choice selection, true/false, gap-fill, etc. The paper then describes the construction, trialling and evaluation of the Text Dictation. In this computer dictation, students hear a recorded text three times. The first time, they simply listen; the second time, they have to type their answers into a text window; the third time, they can click on any of their answers to call the text window back up and to make amendments to what they had previously typed in. Student answers are evaluated on the basis of a parse of the input. Results from trialling the program with different groups of students indicate that the Text Dictation procedure discriminates well between students of differing ability, with significant correlations obtained between students who took the computer dictation test and students who completed a traditional pen-and-paper dictation.

**99–140 Green, Anthony** (Kanda Inst. of Foreign Langs.). Introducing competency-based assessment in a curriculum innovation. *Prospect* (Macquarie U., Sydney), **13**, 1 (1998), 24–39.

Broad-based curriculum innovation must of necessity include development of learner assessment systems. The purpose of this paper is to describe the development of a competency-based assessment system for classroom use in a tertiary vocational college in Japan, which drew on Australian experiences of competency-based training in introducing curricular reform for all its programmes. Competencies are usually expressed as 'can-do' statements which may be used both to guide the course of instruction and to report outcomes; and the aims of competency-based training include ending the sharp traditional distinction between skilled and unskilled workers by increasing access to training, encouraging life-long learning patterns, and thereby creating a flexible, multi-skilled workforce. The paper first reviews the literature relating to competency-based systems in English as both Foreign and Second Language (EFL/ESL). This review is then used as a context for describing the development, trialling and preliminary implementation of the system in this curricular innovation process.

**99–141 Ikeda, Ken** (Nagano Nat. Coll. of Technology, Japan). The paired learner interview: a preliminary investigation applying Vygotskian insights. *Language, Culture and Curriculum* (Clevedon, UK), **11**, 1 (1998), 71–96.

The paired learner interview in which learners take both the role of interviewer and interviewee has been used as a variation of the oral proficiency interview since the 1980s. This paper aims to show that it is an effective means to reduce communicative stress for the non-native learner and one which can also elicit authentic learner participation. Although the oral proficiency interview has come under criticism for lacking authenticity, focusing exclusively on learner performance or proficiency, and preserving assymetrical relations between interviewer and interviewee, the author suggests that the paired learner interview can overcome these cited criticisms if oral proficiency of individual learners is not the only criterion for evaluation, and the interview is structured so that learners can control the interaction themselves. With insights drawn from Vygotskian research, the author suggests that negotiation of meaning, intersubjectivity and scaffolding are critical variables of learner interaction for further study. These variables were investigated in paired interviews of five pairs of Japanese high school students for an end-of-term examination. The paper closes with suggestions on the potential applications of these variables towards assessment and language teaching of the paired learner interview. Analysis of interaction between learners which takes into account the above variables could help ascertain the level of proficiency more comprehensively and accurately with learners from cultures in which speech is held in lower esteem or not encouraged to a great degree. Furthermore, learners in test situations may be motivated to take the initiative to learn new knowledge and incorporate it into their respective private worlds.



https://doi.org/10.1017/S0261444800013628 Published online by Cambridge University Press

**99–142 Johnson, Yuki** (U. of Michigan, USA). Proficiency guidelines and language curriculum making ACTFL Proficiency Guidelines effective in furthering Japanese language proficiency. *Applied Language Learning* (Monterey, CA), **8**, 2 (1997), 261–84.

This article examines the role of ACTFL (American Council on the Teaching of Foreign Languages) Proficiency Guidelines as a system of language proficiency testing, and discusses how effective that system is in relation to university-level classroom instruction. Two main issues are addressed in regard to adapting guidelines into language curriculum. The first is viewing ACTFL Proficiency Tests as having the characteristics of achievement tests. Once the ACTFL Guidelines are adapted into a specific language curriculum, this problem must be addressed since the guidelines cease to function as proficiency testing. The author suggests that, in order to maximise the results from the relationship between proficiency tests and achievement tests, curriculum should be modified so as to embrace a proficiency-oriented instruction which aims to cultivate learners' communicative competence. The resulting type of testing is here termed the 'achievement-proficiency' test. The second main issue is the ACTFL Guidelines' rating scale, which, it is suggested, does not properly reflect major components of communicative competence, even though the goal of testing is to measure this competence. An alternative version is proposed which incorporates theoretical frameworks suggested by Bachman and others and which builds on a strong relationship to classroom instruction. It is concluded that, when the various theoretical and practical issues introduced here are taken into consideration, the ACTFL Guidelines, in their unavoidable guise as achievement-proficiency tests, should result in greater reliability and validity than general proficiency tests and should serve as a helpful tool to aid the process of language acquisition.

**99–143 Lennon, Paul** (U. of Giessen, Germany). Learner-centred testing: a role for cloze? *Language Learning Journal* (Rugby, UK), **17** (1998), 33–40.

This article argues that cloze testing techniques are more learner-centred than others, in particular because cloze is very flexible, but that they also meet the evaluation criteria of validity, reliability, scorability, economy and administrability. Ways are suggested in which cloze can be adapted to particular requirements–e.g., graded for difficulty, constructed so as to test different aspects of language (lexis, syntax, cohesion, etc.); and the pros and cons of rational versus random deletion are discussed, including a comparison of fixed-ratio cloze, rational cloze, C-test, and multiple-choice cloze. Factors requiring consideration in the construction of a cloze test are also considered. The article also examines several general aspects of language testing: the function of different sorts of tests (diagnostic, classroom achievement, proficiency, etc.), evaluation criteria, the tendency of tests to distract both teachers and pupils from the main activity of language teaching and learning, the drawbacks of integrated and discrete point tests, and how the motivation of testees is affected by giving them relevant and accessible feedback.

**99–144 Lynch, Brian K. and McNamara, T. F.** (U. of Melbourne, Australia). Using G-theory and Many-facet Rasch measurement in the development of performance assessments of the ESL speaking skills of immigrants. *Language Testing* (London), **15**, 2 (1998), 158–80.

Second language performance tests, through the richness of the assessment context, introduce a range of facets which may influence the chances of success of a candidate on the test. The study reported here investigates the potential roles of Generalisability theory (G-theory) and Many-facet Rasch measurement in the development of such a performance-based assessment procedure. This represents an extension of preliminary investigations into the relative contributions of these procedures to another assessment setting. Data for the present study come from a trial of materials from the access test, a test of communicative skills in English as a Second Language (ESL) for intending immigrants to Australia. The performances of 83 candidates on the speaking skills module were multiply rated and analysed using GENOVA and FACETS. The advantages and specific roles of these contrasting analytical techniques are considered in detail in the light of this assessment context.

**99–145 North, Brian** (Eurocentres Foundation, Zürich, Switzerland) **and Schneider, Günther**. Scaling descriptors for language proficiency scales. *Language Testing* (London), **15**, 2 (1998), 217–63.

This paper reports results from a Swiss National Science Research Council project which aimed to develop a scale of language proficiency in the form of a 'descriptor bank'. Up until now, most scales of language proficiency have been produced by appeal to intuition and to those scales which already exist rather than to theories of linguistic description or of measurement. The intention in this project was to use an item-banking methodology to develop a flexible scale of stand-alone criterion statements with known difficulty values. The project took place in two rounds: the first for English (1994), the second for French, German and English (1995). In each year pools of descriptors were produced by analysing available proficiency scales. Through workshops with representative teachers, the descriptors were then refined into stand-alone criterion statements considered to be clear, useful and relevant to the sectors concerned. Selected descriptors presented on questionnaires were then used by participating teachers to assess the proficiency of learners in their classes. These data were used to scale the descriptors using the Rasch rating scale model. The difficulty estimates for the descriptors produced in relation to



https://doi.org/10.1017/S0261444800013628 Published online by Cambridge University Press

English in 1994 proved remarkably stable in relation to French, German and English in 1995.

**99–146 Paulus, Linda** (Mundelein High Sch., USA). Watch them SOAR: Student Oral Assessment Redefined. *Hispania* (Ann Arbor, MI), **81**, 1 (1998), 146–52.

This article describes an approach to oral assessment at the first year high school level which is claimed to result in a strong sense of student and teacher satisfaction. The approach is based on a model of communicative competence comprising four components–Strategic, Sociolinguistic, Discourse and Linguistic Competencies–and consists of evaluating students' oral performance in these different domains. It is suggested that this competence assessment model drives instruction in a way that facilitates the creation of an immersion classroom and enables students to acquire language in a more meaningful way than by the use of models in which students are globally evaluated, or evaluated on linguistic competence alone.

**99–147 Sparks, Richard L.** (Coll. of Mount St. Joseph, Cincinnati, OH, USA), **Artzer, Marjorie, Ganschow, Leonore, Siebenhar, David, Plageman, Mark and Patton, Jon.** Differences in native-language skills, foreign-language aptitude, and foreign-language grades among high-, average-, and low-proficiency foreign-language learners: two studies. *Language Testing* (London), **15**, 2 (1998), 181–216.

This paper reports two studies that examined the extent to which there would be differences in native language skills, foreign-language aptitude and final foreign-language grades among high-school students completing a second year of a foreign-language course and identified as high-, average- and low-proficiency learners. Oral and written proficiency measures in Spanish, French and German were designed and administered by trained evaluators. The first study involved 60 females attending a private, single-sex suburban high school; the second involved a co-educational population of 36 students in a suburban public (i.e., state) school. Results showed overall differences among the three proficiency groups on native-language and foreign-language aptitude measures. Most group differences were between high- and low-proficiency learners, although high *vs.* average and average *vs.* low groups differed on some measures. There were also group differences in final grades at end of both first-year and second-year foreign-language courses. Discriminant analyses showed that two testing measures in the first study and one in the second study were significant in distinguishing the three groups. Conclusions and implications about connections among foreign-language proficiency and native-language skills, foreign-language aptitude, and end-of-year grades are presented.

**99–148 Tschirner, Erwin and Heilenman, L. Kathy** (U. of Iowa, USA). Reasonable expectations: oral proficiency goals for intermediate-level students of German. *The Modern Language Journal* (Madison, WI), **82**, 2 (1998), 147–58.

As an increasing number of American states institute oral proficiency exit requirements for high school students, and higher numbers of universities and colleges establish entrance and outcome requirements including an oral proficiency component, it becomes imperative that the profession gain a clear sense of what proficiency levels are possible given the amount of time and type of instruction commonly available at high school and college level. This article reports a study which set out to gauge the average oral proficiency of college students ($N$ = 20) studying German at the end of their fourth semester. The measure used was the ACTFL (American Council on the Teaching of Foreign Languages) Oral Proficiency Interview (OPI). All OPIs were official ACTFL double-rated interviews. Proficiency, as measured by the OPI, ranged between Novice High and Intermediate Mid. The discussion here centres on reasonable expectations for high school graduation and for college language requirements.

**99–149 Weigle, Sara Cushing** (Georgia State U., USA). Using FACETS to model rater training effects. *Language Testing* (London), **15**, 2 (1998), 263–87.

This article describes a study conducted to explore differences in rater severity and consistency among inexperienced and experienced raters both before and after rater training. Sixteen raters–eight experienced, eight inexperienced–rated overlapping subsets of essays from a total sample of 60 essays before and after rater training in the context of an operational administration of the English as a Second Language Placement Examination (ESLPE) of UCLA (University of California, Los Angeles). A three-part scale was used, comprising content, rhetorical control and language. Ratings were analysed using FACETS, a multi-faceted Rasch analysis programme that provides estimates of rater severity on a linear scale as well as fit statistics, which are indicators of rater consistency. The analysis showed that the inexperienced raters tended to be both more severe and less consistent in their ratings than the experienced raters before training. After training, the differences between the two groups were less pronounced; however, significant differences in severity were still found among raters, although consistency had improved for most of them. These results provide support for the notion that rater training is more successful in helping raters give more predictable scores–i.e., intra-rater reliability–than in getting them to give identical scores–i.e., inter-rater reliability.

# Teacher education

**99–150 Cadorath, Jill** (U. of Yucatan, Mexico) **and Harris, Simon.** Unplanned classroom language



https://doi.org/10.1017/S0261444800013628 Published online by Cambridge University Press