Assessment of Business and Professional Language for Specific Purposes

JANE LOCKWOOD

Introduction

The use of English in international worksites and across business and professions has grown rapidly since the 1980s. Multinational companies doing business worldwide need to communicate and negotiate with each other; global professionals need to build knowledge together across linguistic and cultural boundaries; governments and NGOs need to keep the peace and help build developing countries; and immigrants with professional and occupational aspirations and qualifications need to find work in their new homelands. These situations predominantly use English as their lingua franca (Graddol, 2006). English-language competency has become a global core skill and a highly valued commodity in today's world of work. Therefore developing and measuring business and professional language competence is understandably in high demand.

This entry will attempt to provide an overview of the key issues in English-language assessment for the specific purposes of business and the professions both now and in the future. The entry starts by describing what sorts of assessments are currently in the marketplace for business and professional English testing and evaluates their strengths and weaknesses. It then goes on to examine some current trends in English for Specific Purposes (ESP) curriculum pedagogy and finally discusses how this thinking may affect language-assessment practices in the future.

Recent Approaches to Language for Specific Purposes (LSP) Assessment in Business and Professional Contexts

Since the 1980s, and in response primarily to the business and professional needs of prospective employees from non-English-speaking backgrounds, LSP tests have proliferated. The purpose of these tests have, in the main, been to provide governments, educational providers, professional associations, and workplaces with valid, reliable, and practical tools to screen for the language suitability of the LSP test taker for, among other things, employment, promotion, citizenship, and university entrance. They have been developed to provide a one-point-in-time "gatekeeping" benchmarked screening test. These high-stakes screening LSP assessments fall into two broad categories.

First there are the large-scale "specific" language assessments developed by large testing organizations such as Cambridge ESOL, Educational Testing Services, and Pearson Education. In turn, these assessments fall into two main specialisms: English for tertiary studies and English for business. Over the years, such tests as the International English Language Testing System (IELTS; Cambridge ESOL) have developed a pre-tertiary test of English for non-English-speaking-background (NESB) high-school leavers wanting to gain entrance into English-speaking universities. The Test of English as a Foreign Language (TOEFL; ETS) and the Pearson Test of Academic English (PTE; Pearson) have been developed for this same large pre-university international market.

For the business market, these same large testing organizations, with the exception of Pearson, have developed the Business Language Testing Service (BULATS), the Business

English Certificate 1–3 (BEC)—both offered by Cambridge ESOL—and the Test of English for International Communication (TOEIC)—offered by ETS—to name only a few. All these tests, however, face significant development challenges to ensure they maximize their market share of candidates who may have many different reasons for sitting them. For example, the test takers of the business tests are not necessarily interested in using their language-test result to work in a Western English-speaking country or even in a Western English-speaking company. For example, a university business graduate from Korea may want a TOEIC result to do business in Asia where the lingua franca is English. He or she may be communicating most of the time to non-native speakers of English. Because of the need for these business-testing organizations to maximize their market share, they have to ensure that they are not promoting specific business or ethnic cultures, nor specific language varieties and accents within their rubrics and test tasks. For example, asking a Japanese woman in a speaking assessment to recount a time when she lost her temper in a work situation is linguistically and culturally inappropriate, as may be role-play topics that too readily assume aligned views on such things as work–life balances, good parenting, the status of women as managers, and the responsibilities of the wealthy. In one well-known business-English test being administered in Hong Kong, the test task was seriously compromised because the notion of a "conference center" in the UK sense of the term was not readily understood in the Hong Kong context. These large assessment agencies are constantly modifying their tests and to some extent "neutralizing" the task content so that it does not favor or disadvantage any particular LSP business-English test-taking group.

The major benefit of these business LSP tests is that they provide an instantly recognizable international benchmark for language-proficiency level with validity and reliability. This is evidenced by the large number of universities and workplaces that stipulate ETS (TOEFL or TOEIC) or Cambridge ESOL (BULATS) benchmark levels for entry. These tests are what could be termed "generic" LSP tests for business and the workplace.

The other category of more genuinely "specific" language assessments has generally been developed by universities, professional bodies, and private companies, mostly in English-speaking countries in the West, to meet the needs of specific professional groups. The word "genuinely" is used here because these tests have been developed often at the request of clients to test highly specific communication skills for employment. These tests therefore attempt to mirror the communication needs of very specific target-language-use (TLU) situations of a range of professional workers.

These professional entry tests differ from the first category discussed in this entry in two ways. First, they are used for smaller and homogeneous professional and occupational populations; and, second, they use specialized and contextualized knowledge and authentic data taken from an analysis of the TLU situation to inform the test task and rubric development.

> A specific purpose language test is one in which test content and methods are derived from an analysis of a specific purpose target language use situation, so that the test tasks and content are authentically representative of tasks in the target situation, allowing for an interaction between the test taker's language ability and specific purposes content knowledge on the one hand, and the test tasks on the other. Such a test allows us to make inferences about a test taker's capacity to use language in the specific purpose domain.
> (Douglas 2000: 19)

It is noteworthy that many of these tests have been developed by countries such as the USA, Canada, and Australia—countries with very high immigrant populations seeking work. For example, doctors and nurses in Australia need to pass the Occupational English

Test (OET) funded by the Australian government and endorsed by the Australian Medical Association (AMA) in order to become registered health professionals in Australia. McNamara (1990, 1997) has written extensively about the development, implementation, and evaluation of the OET. This test was developed as the result of exhaustive observations of the TLU situation and an analysis of authentic recordings of doctor–patient exchanges. Similarly, school teachers have to pass the Proficiency Test for Language Teachers (McDowell, 1995; Elder, 2001), also funded by the Australian government, to become classroom teachers in Australia. Airline pilots worldwide, as of 2008, need to sit a screening test from a range of English for aviation tests as mandated by the International Civil Aviation Organization (ICAO) to have a certificate attesting to their proficiency level in the language(s) used for aeronautical communication. Even peacekeepers are tested for English before they go on assignments overseas where the common language will be English. In the development of some of these tests, new and innovative linguistic frameworks have been used in linking the test task to the target-language context through using genre studies, discourse analysis, close TLU observations, and corpus studies to inform the test. In this kind of LSP assessment development the immediate language-test-use situation, analyses of the authentic data, and the veracity of the test tools and processes are paramount. However, is this enough? Are these the key requirements for LSP assessment?

Some Issues to Ponder

It is often the case, as reported by Douglas (2001) and Jacoby and McNamara (1999), that, despite passing the specialized test, the candidate still fails to communicate in an appropriate way on the job. What more can be done by the assessment designer to ensure that communication weaknesses are identified and the criteria revisited?

Perhaps there is a need to consider in much more detail the professional contexts, purposes, and stakeholder practices for which these tests are requested. However, to do this, LSP developers need access to the business and professional stakeholder concerns to explore how the requirements of their own (professional stakeholder) evaluation processes can be captured through language assessments. An emerging literature in ESP curriculum development, suggesting that ESP practitioners consider and understand more deeply the concerns of the workplace, is setting a research agenda also for those working in LSP assessment development. Let us first consider the current work in ESP curriculum development as a way into these concerns.

Since the 1980s there has been a proliferation of ESP courses ranging from English for health professionals to English for call-center agents, and the curriculum-development issues are well documented (Hutchinson and Waters, 1987; Johns & Dudley-Evans, 1993; Candlin & Candlin, 2003; Basturkman, 2006; Bremner, 2008; Belcher, 2009; Bhatia, 2008; Lockwood, 2010). Genre-based pedagogies (i.e., the way spoken and written texts are organized), discourse analysis (i.e., the way certain lexico-grammatical linguistic features are reflected and patterned in the text under investigation), and corpus linguistics (i.e., the frequency of occurrence of certain lexico-grammatical items found in the authentic data of the texts) have provided new and exciting linguistic frameworks to probe the specialization of language knowledge, skills, and awareness required in specific worksites and occupational-language texts. This current theory is informing new and innovative curriculum practices. Also, interdiscursivity approaches (i.e., the way texts in professional sites will build in certain ways on each other; Bremner, 2008; Bhatia, 2008) in the work of ESP specialists are currently being suggested as a means for facilitating the understanding of learners' communication needs in ESP courses. These linguists talk about interdiscursivity and the importance of multiple perspectives for ESP curriculum design, calling for

much more than a linguistic analysis of professional genres: "Professional genres have often been analysed in isolation, leaving the study of professional practice almost completely out, except as providing context for specific analyses, thus undermining the role of inter-discursivity in professional genres and practices" (Bhatia, 2008, p. 161). Bhatia defines this concern further when he says in a more recent article:

> Other contributions to the construction of professional artefacts are conventions of the genre in question, the understanding of the professional practice in which the genre is embedded, and the culture of the profession, discipline, or institution, which constrains the use of textual resources for a particular discursive practice. In other words, any instance of professional communication simultaneously operates and can be analysed at the four levels, as text, as representation of genre, as realization of professional practice, and as expectation of professional culture. (Bhatia, 2010, p. 33)

An empirical exploration of the TLU contexts is being suggested in these approaches.

LSP assessment practices appear to have lagged behind the innovations in course development where communicative competency domains (Canale & Swain, 1980; Bachman & Palmer, 1996) have dominated LSP assessment practices across the four skills of listening, speaking, reading, and writing. Douglas (2001) challenges these criteria and has written more recently about what he terms "indigenous criteria" for LSP assessment:

> I want to discuss the theoretical underpinnings of the development of LSP assessment criteria and argue that it is important for us to derive them not only from a theoretical understanding of communicative language ability (Bachman and Palmer, 1996) but also from an empirical analysis of the TLU situation. (Douglas, 2001, p. 174)

Douglas's concern was triggered by reports from subject specialists at his own university that students were judged on criteria other than communicative-competency domains when delivering their presentations, and he was interested to find out why this was the case. How deeply one should conduct an empirical analysis of the TLU situation is an interesting question and Douglas concludes:

> There is a "strong" indigenous assessment hypothesis which would involve employing the criteria derived from an analysis of assessment practices directly in the TLU situation; however, I do not advocate such a strong case. Rather I wish to suggest a weaker indigenous assessment hypothesis in which the indigenous criteria may be used to supplement linguistically-oriented criteria in line with construct definition, and, secondly, to help guide our interpretations of language performances in specific purpose tests. (2006, p. 183)

This is indeed a challenge to the LSP assessment fraternity in selecting, validating, and calibrating new criteria.

Some Challenges Ahead

LSP assessment is an emerging research and development area and provides the testing community with challenges. These are discussed in this part of the entry. First, we must ask the question about how we ensure that LSP assessment development is context sensitive and to what degree. How does the assessment meet the changing needs within the requirements of the industry, worksite, or educational institution? Are there arguments for using what Douglas (2001) has termed "weak" indigenous assessment criteria as opposed

to "strong" criteria in certain contexts? If indigenous criteria for assessments become written into LSP assessments, how is this going to affect the validity and reliability of our assessments?

This then leads to the question of "ownership" of the LSP assessments. Who decides ultimately on the criteria? Coupled with this is the question about who carries out LSP assessments. Conventional practice in LSP testing is that the assessments themselves are carried out by applied-linguistic experts suitably trained as language assessors. However, workplaces are challenging this practice. Why can quality-assurance personnel in banks not assess communication skills? What are the issues in training such personnel to carry out LSP assessments?

Also, given that some workplaces are requiring LSP language assessment to be embedded in their own work practices (Lockwood, 2010), is there an LSP assessment process that might be developed to ensure that ongoing workplace LSP assessment happens in a systematic and thoroughgoing way that remains valid and reliable and embedded in the work flow of that worksite? This would necessitate close stakeholder involvement and ownership and a thorough understanding of the contextual requirements for assessment to be incorporated, end-to-end, into workplace practices.

Another significant challenge for LPS assessment relates to the fact that traditional paradigms for LSP assessment development are also being challenged by the growing acceptance that native-speaker norms are no longer relevant to many (if not most) target-language LSP assessment contexts, especially in international business contexts. So much of the LSP assessment work to date has assumed a native-speaker model. In a world where most of the business communication is conducted between NESB speakers (Graddol, 2006), how can this be relevant and fair? How can the concerns of business English as lingua franca (BELF) be reflected in LSP assessment tools and processes in the future (Nickerson, 2005)?

One final concern for LSP assessment developers is the situation where LSP assessments are used for contexts for which they were not designed. For example, the IELTS academic-paper exam is being used for high-stakes professional employment purposes and for university-graduation requirements in many countries. TOEIC scores for listening and reading are being used to recruit customer-services agents in the burgeoning business-processing outsourcing (BPO) industry in some parts of Asia. How can highly lucrative LSP assessment contracts be regulated for the validity and reliability of the tests they offer?

In conclusion, it would seem that some of the greatest challenges for LSP assessment lie ahead. Gaining access to LSP assessment worksites, listening to and learning about their needs, and keeping an open mind on who needs this assessment expertise and how it can be sensibly transferred and embedded in the workflow of the company are critical.

The need to offer valid, reliable, and practical tailor-made solutions will provide a busy agenda for LSP assessment in the future for businesses and professions.

SEE ALSO: Assessment of Listening; Assessment of Speaking; Assessment of Writing; Douglas, Dan; English as Lingua Franca; English for Business; English for Occupational Purposes; Task-Based Assessment

References

Bachman, L., & Palmer, A. (1996). *Language testing in practice*. Oxford, England: Oxford University Press.
Basturkman, H. (2006). *Ideas and options in English for specific purposes*. Mahwah, NJ: Erlbaum.

Belcher, D. (Ed.). (2009). *English for specific purposes in theory and practice*. Ann Arbor: University of Michigan Press.

Bhatia, V. (2008). Genre analysis: ESP and professional practice. *ESP Journal, 27*, 161–74.

Bhatia, V. (2010). Interdiscursivity in professional communication. *Discourse & Communication, 4*(1), 32–50.

Bremner, S. (2008). Intertextuality and business communication textbooks: Why students need more textual support. *English for Specific Purposes, 27*, 306–21.

Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics, 1*, 1–47.

Candlin, C. N., & Candlin, S. (2003). Healthcare communication: A problematic site for applied linguistic research. *Annual Review of Applied Linguistics, 23*, 134–54.

Douglas, D. (2000). *Assessing languages for specific purposes*. Cambridge, England: Cambridge University Press.

Douglas, D. (2001). Language for specific purposes criteria: Where do they come from? *Language testing, 18*, 171–85.

Elder, C. (2001). Assessing the language proficiency of teachers: Are there any border controls? *Language Testing, 18*(1), 149–70.

Graddol, D. (2006). *English next*. London, England: British Council.

Hutchinson, T., & Waters, A. (1987). *English for specific purposes: A learner-centred approach*. Cambridge, England: Cambridge University Press.

Jacoby, S., & McNamara, T. (1999). Locating competence. *English for Specific Purposes, 18*(3), 213–41.

Johns, A. M., & Dudley-Evans, T. (1993). English for specific purposes: International in scope, specific in purpose. *TESOL Quarterly, 25*, 297–314.

Lockwood, J. (2002). *Evaluating training in Hong Kong workplaces* (Unpublished doctoral dissertation). Hong Kong University.

Lockwood, J. (2010). Consulting assessment for the business processing outsourcing (BPO) industry in the Philippines. In G. Forey & J. Lockwood (Eds.), *Globalization, communication and the workplace* (pp. 221–41). London, England: Continuum.

McDowell, C. (1995). Assessing the language proficiency of overseas-qualified teachers: The English language skills assessment (ELSA). In G. Brindley (Ed.), *Language assessment in action* (pp. 11–29). Sydney, Australia: NCELTR.

McNamara, T. (1990). *Assessing the second language proficiency of health professionals* (Unpublished doctoral dissertation). University of Melbourne.

McNamara, T. (1997). Interaction in second language performance assessment: Whose performance? *Applied Linguistics, 8*, 446–66.

Nickerson, C. (2005). English as a lingua franca in international business contexts. *English for Specific Purposes, 24*(4), 367–80.

Suggested Readings

Alderson, J. C. (2009). Air safety, language assessment policy and policy implementation. *Annual Review of Applied Linguistics, 29*, 168–87.

Bachman, L. (1990). *Fundamental considerations in language testing*. Oxford, England: Oxford University Press.

Davies, A. (2009). Assessing World Englishes. *Annual Review of Applied Linguistics, 29*, 80–9.

Hamp-Lyons, L., & Lockwood, J. (2009). The workplace, the society, and the wider world: The offshoring and outsourcing industry. *Annual Review of Applied Linguistics, 29*, 145–67.

McNamara, T., & Roever, C. (2006). *Language testing: The social dimension*. Malden, MA: Blackwell.