

RESEARCH

Open Access

Language for Specific Purpose (LSP) performance assessment in Asian call centres: strong and weak definitions

Jane E Lockwood

Correspondence:

lockwood@cityu.edu.hk
Department of English City,
University of Hong Kong, 18, Tat
Hong Rd, Kowloon Tong, Hong Kong

Abstract

LSP performance assessment has become a special focus for language testers in recent years where experts have debated how testing tools and processes can be strengthened to more accurately and more validly assess professional communication at work. Suggestions to achieve this include ethnographic studies of the target language situation; authentic discourse analyses of the relevant texts; and subject matter experts (SMEs) being invited as informants when defining 'successful' communication at work. These are proposed as key steps in building LSP performance assessment validity (see for example Jacoby & McNamara: *English for Specific Purposes* 18(3):203–241, 1999). Language testing researchers have also suggested there may be some merit in distinguishing LSP performance as 'weak' and 'strong' (see also McNamara: *Measuring Second Language Performance*, 1996; Douglas: *Language Testing* 18(2):171–185, 2001) although clear lines of distinction between these have not yet been made in LSP assessment studies to date.

In this article I propose a distinction between a 'weak' and 'strong' version of LSP performance assessment based on the empirical data collected over the last ten years of developing and embedding the Business Performance Language Assessment Scales (BUPLAS) into Asian call centres. A distinction that has been unclear to date in this area of research is whether a workplace LSP spoken assessment for work (meaning those assessment tasks that gauge employment entry levels as 'predictive' of work success) is the same as a workplace LSP spoken assessments at work (meaning those authentic assessments that gauge quality levels on-the-job as 'observed' success at work). For the Asian call centre industry this distinction is important because the purpose of recruitment (exclusion) is very different from the purpose of quality assurance (appraisal and coaching feedback). I argue that recruitment assessment with its attendant characteristics constitutes a 'weak' version of LSP performance assessment whilst its quality assurance counterpart constitutes a 'strong' version.

Keywords: Language for specific purposes; Assessment; Communication at work

Introduction

The outsourced call centre industry in India and the Philippines considers excellent English language communication skills to be its core commodity for business success. This is because the call centre customer services representatives (CSRs) have to meet high quality communication and business targets when serving Western English speaking customers from around the world. Recruiting, training, appraising and coaching

for excellent levels of English in this industry therefore require assessment tools and processes to fit the business requirements and demonstrate impact. Unfortunately only a small percentage of applicants gain employment in this industry (see The Business Processing Association of the Philippines BPAP Report 2006) due to inappropriate 'gate-keeping' entry assessment practices (see Lockwood, 2012). However, the industry invests heavily in training and coaching their CRSs who are meticulously tracked for communication 'performance' on the phones on a regular basis. The work of the CRSs involves the delicate balancing of profiling the customer, understanding the request, managing the multiple screens locating product and customer information, and responding professionally and politely within average handling time (AHT) limits. Quality assurance processes are key to the success of the company and these quality measures are used as an important measure in contractual and pricing agreements; they are therefore scrutinized closely and regularly by on-shore management. It is therefore not surprising that being able to measure this highly specialized form of spoken performance validly and reliably has emerged as a significant business imperative in these call centres.

LSP performance assessment in professional and educational contexts has been the subject of applied linguistic study and debate over the last two decades where researchers have been working in such high stakes contexts as aviation safety, healthcare and education (see for example Elder, 2001; Elder et al. 2012; Alderson, 2011). This paper positions call centre LSP performance assessment within this debate and within an Asian business setting. The Business Performance Assessment System (BUPLAS) instrument is used as the basis for this study; I am not a disinterested party in BUPLAS as this assessment approach was first conceived and developed by me when I was working in the Philippines (2003–2005) as communications consultant; but has, since then been further developed by language specialists employed by a communications training company for the call centre industry (called Company A for the purposes of this article). Data for this journal article has been supplied by Company A.

Building on existing labellings of 'weak' and 'strong' LSP assessment (see for example McNamara, 1996; Douglas, 2001), this study proposes clear definitions for each of these where BUPLAS is used for recruitment and quality assurance in the Asian call centre context.

I argue that the essential differences between the 'weak' and 'strong' versions of BUPLAS are characterized in their specific purposes and impact value; in their locations; in their raters; in their tasks, and ultimately, in the ways in which the common BUPLAS criteria are ranked, weighted and elaborated. For example, BUPLAS used for recruitment (the 'weak' form) is characterized by drawing on recruiter rater inferences and predictions for employment success through assessing on a phone interview; whereas BUPLAS for appraisal and coaching feedback (the 'strong' form), used by subject matter experts, is characterised by assessing direct observations of live or recorded authentic call exchanges. I explore how LSP assessment in each of these recruitment and quality assurance business points is carried out and how the specific needs in these two points impact the number, the selection, the description and the rankings and weightings of the criteria used for LSP spoken performance assessment in this industry. I further explore how the tasks and the raters also differ in each site, ultimately leading to enhanced validity for LSP performance assessments in their 'weak' and 'strong' forms in this specific Asian business context.

Literature review

Research and development in LSP performance assessment is a relatively new field of enquiry and is particularly focused on the high stakes areas of global professional and business communication performance in the assessment of teachers, medical personnel, air traffic controllers and pilots (Elder, 2001; Elder et al., 2012; Alderson, 2011). These studies seek to answer the question increasingly posed by globalized workplaces as to why employees with good scores in spoken proficiency on such generic international English language tests, for example the International English Language testing System (IELTS) and business English tests such as the Test of English for International Communication (TOEIC) often fail to perform well at work. It would appear that spoken language performances in specific contexts, for example in patient interaction, air traffic controller/pilot exchanges, academic conference presentations; and in this case, call centre exchanges, require a good understanding of that context, a good understanding and analysis of the target texts in that context, and a reliance on subject matter expert (SME) input as to what constitutes 'task fulfillment' to achieve validity in that context. Using the concept of 'indigenous criteria', meaning what occupational/professional stakeholders value in communicative exchange at work, (see Jacoby, 1998; Jacoby and McNamara, 1999; Douglas, 2001, 2005), language testers have looked at different ways and sources from which these criteria may be derived. McNamara (1996) and Jacoby and McNamara (1999), first became interested in how SMEs and language assessors (LAs) differed in their evaluations of successful spoken performance:

"since members of professional cultures are faced with the interactional problem of articulating to one another the grounds for their own evaluations of performance and recommendation for improvement, an analysis of 'indigenous assessment' may provide a window onto an insider's view of the complex issues involved in communicating competently in some particular domain" (Jacoby and McNamara 1999: 214).

Douglas (2000, 2001) and Hamp-Lyons and Lumley (2001) later argued that the successful sourcing of these professionally relevant criteria to be one of the most urgent needs for research and practice in LSP assessment.

McNamara (1996) proposed 'the use of the terms *strong and weak second language performance test* to reflect the extent to which the assessment criteria reflect real-world contexts, or whether they are oriented mainly to linguistic aspects of the communication respectively' (p. 217). He also suggested that an investigation of 'naturalized assessment practices in everyday and in professional life' (p.43) would be helpful in understanding how LSP assessment practices might work. This distinction is highly relevant to the current study which seeks to differentiate the characteristics of 'weak' and 'strong' LSP spoken performance assessment in the call centre workplace; this study however pushes this distinction beyond just a consideration of the criteria as originally suggested by McNamara (1996).

To date there have been only limited studies into spoken language assessment approaches in the call centre context (see for example, Lockwood et al., 2009, Lockwood 2010, Lockwood 2012; Davies 2010; Friginal 2013). This is unsurprising given that outsourced contact centres have only been operating outside western English countries for a little over a decade and thus the need for assessing second language English speaker CSRs is a relatively new business requirement in this rapidly developing global industry. Previous studies in call centres (see for example Forey and Lockwood 2007;

Lockwood et al., 2008) had revealed however that the source of communication breakdown on the phones 'has less to do with the traditional notions of poor language skills i.e. poor grammatical use and poor pronunciation, and more to do with poor interaction and discourse skills and cultural appreciation' (p.323).

The call centre context

Typically in this industry, spoken assessment tasks and rating scales have been developed internally by the business and/or have been sourced off the internet; but without a good applied linguistic understanding about what they should contain and how they should be administered, they have, in general, proven to be unreliable and invalid (Lockwood 2010, Lockwood 2012). For example, many CSR applicants are often rejected at recruitment because the business managers worry about the narrow concerns of grammatical mistakes and accents, citing that they 'sound too Filipino', 'they don't have the right accent' or that 'they make too many grammar mistakes'. This limited view of language has had a negative washback into internal assessment tools developed for this industry. 'Accent' and 'grammar accuracy' unfortunately persist as the key indicators of functional language ability in the minds of many call centre managers today. Quality assurance appraisal on-the-floor is measured against a tool called a 'scorecard' that similarly privileges accent and grammar accuracy (see Lockwood et al. 2009 for a full discussion about scorecard use in call centres). These quality scorecards ideally aim to assess business product and process knowledge and communication skills during live or recorded authentic calls for judgmental (appraisal) and developmental (coaching) purposes. Developing a 'strong' LSP performance measure, as informed by applied linguistic theory and practice, and involving ethnographic studies, analyses of the authentic exchanges and SME input into a revised scorecard, is argued to add validity to this tool.

This article contributes to the on-going debate defining the nature of LSP performance in its different forms by arguing that 'strong' performance assessment validity relies on the incorporation of business specific 'indigenous criteria' as informed by SMEs; and furthermore it is assessed through the SME rating of authentic calls thus reflecting the real-world context of the call centre. Weak LSP performance validity on the other hand, could be said to be more linguistically (and communicatively) oriented fulfilling the task of employment entry benchmarking and rated by communications trainers and recruiters. These characteristics will be further elaborated in this article.

BUPLAS has been used as the basis for this study and I describe the background to its development in the next section.

The business processing language assessment system (BUPLAS) assessment

A typical call centre employs teams of CSRs who work for a specific account and deal with a specific product and/or process. For example a big credit card company may have about 100 CSRs working in its Collections Account where global customers are asked to pay their overdue credit card bills. Each account is managed by an Account Manager and teams of 12 CSRs are managed day-to-day by the next level down of management, the Team Leads. Business targets are set and monitored, as are the calls. Metrics measuring instances of product promotions and compliance as well as first time resolution (FTR) and average handling time (AHT) are tabulated daily. Quality

Assurance Specialists (QASs), who may also be Account Managers or Team Leads make monthly assessments for appraisal and coaching support purposes. The QAS is therefore an SME, and this role is often a part-time one and combined with other business management and operational roles and responsibilities in the call centre.

The BUPLAS assessment tool was first developed by Company A, a business English consultancy firm based in Manila over a decade ago, and has gradually been refined, in close collaboration with the business, to meet different purposes within the call centre workplace (Lockwood 2010). These call centre assessment purposes are divided into two: first BUPLAS for recruitment purposes, and secondly, BUPLAS for quality assurance purposes. As suggested earlier, I will argue that BUPLAS recruitment (known by the industry users as BUPLAS Voice Assessment (VA)) is an example of a 'weak' LSP spoken assessment and BUPLAS quality assurance (known as BUPLAS Call Assessment (CA)) is an example of a 'strong' version. I will deconstruct the characteristics of each later in this article in support of these proposed definitions.

The BUPLAS system is provided on a licensed basis and has achieved successful penetration into outsourced call centres in the Philippines and India where it is estimated that over a million contact call candidates have taken the BUPLAS recruitment assessment (see BUPLAS.com).

The BUPLAS construct

The theoretical construct for BUPLAS draws on sociolinguistic frameworks informed by Systemic Functional Linguistics (SFL) (Hasan, 1985; Halliday 1985; Ventola, 2005) and on the early work of Hymes, 1973; Canale and Swain 1980; Canale, 1983; Bachman and Savignon 1986; Bachman, 1990, in an effort to define 'communicative competence' for this workplace. Such sociolinguistic frameworks are ideally suited to the language performance assessment of service encounters where success in the language exchange entails much more than demonstrating a baseline proficiency of spoken grammar and pronunciation ability. Interactional, discourse, intercultural and strategic competencies are key in successful call centre exchange as revealed in previous studies (see for example, Forey and Lockwood 2007; Hood and Forey, 2008). For a fuller discussion on the development of BUPLAS see Lockwood 2010, Lockwood 2012.

The assessment approach embedded into the development of BUPLAS has been a 'knowledge and skills' transfer to the key company stakeholders for their own use. There is a great deal of business management literature on the benefits of the smooth transfer of knowledge and skills into and within organisations where 'knowledge embedded in the interactions of people, and tasks provides a basis for competitive advantage in firms' (Argote and Ingram 2000:150). Typically it is now communications trainers, recruitment personnel, operational managers and quality assurance specialists who are trained and calibrated to use the BUPLAS assessment system. This process of knowledge transfer to SMEs has been a contentious step in that it departs from traditional practices of language communications assessment where a TESOL background is seen as a requirement of a language assessor (see for example the guidelines for assessor qualifications for the Canadian Language Benchmark Assessment (CLBA); the Education Testing Service (ETS) and the International English Language Testing System (IELTS)). However whilst this particular aspect of the BUPLAS assessment is not the key focus of this article, it does contribute to my definitions of 'weak' and

‘strong’ LSP assessment in that a ‘strong’ version of LSP performance assessment for BUPLAS relies on solely on internal SMEs, with no language teaching experience, to work as reliable assessors.

I will first consider the two versions of BUPLAS currently used in the call centres. This will then be followed by a synthesis of the specific features of each resulting in ‘weak’ and ‘strong’ definitions of BUPLAS.

BUPLAS voice assessment (VA) recruitment instrument: a ‘weak’ performance assessment

Recruitment LSP assessment, defined as ‘weak’ LSP assessment is an ‘exclusion’ language proficiency test aimed at employment entry. This assessment is also viewed and valued by the business as predictor of success for call centre work. This LSP performance assessment FOR work, is now known within the industry as BUPLAS voice assessment (VA). The judgments are made against four equally weighted linguistic criteria that are communicatively oriented, as shown in Table 1. The phone interview, lasting 8–10 minutes, comprises three tasks with the rater as interlocutor. The tasks include a ‘read aloud’ passage to check comprehensibility of accent and pronunciation; an extended talk on a prepared topic followed by an in-depth interactive discussion as the rater poses progressively more challenging questions and statements to the candidate for spontaneous response. The rater is a BUPLAS VA trained assessor and will typically have worked as a communications trainer in the call centre. The BUPLAS VA assessor training takes place over 2 days and is followed up by multiple calibration sessions before they are accredited.

The criteria, weightings and focus areas are detailed below.

BUPLAS has five levels and ranges from roughly a Common European Framework of Reference (CEFR) B1(lowest) to a CEFR C2 (highest). Half scores are used on the 5 point scale to achieve granularity in proficiency levels. Different accounts within the call centre may require a higher or lower BUPLAS levels across the domains depending, for example, on the account complexity and/or sensitivity. Ideally candidates will be matched to appropriate accounts according to their BUPLAS VA levels.

BUPLAS call assessment (CA) quality assurance assessment: a ‘strong’ performance measure

In contrast to BUPLAS VA, quality assurance LSP performance assessment, called by the industry, BUPLAS call assessment (CA), comprises the same linguistic criteria but with an added business focused task fulfillment one called ‘Business Solutioning’ (see

Table 1 BUPLAS for recruitment

BUPLAS recruitment criteria Voice Assessment (VA)	Focus
Pronunciation (+ intonation, word stress) -25%	This domain poses the questions as to whether the CSR is globally comprehensible and is using the prosodic features to make meaning appropriately.
Lexico-grammatical choices -25%	This domain poses the questions as to whether the choice in the grammar and vocabulary is of a sufficient range for the communication task at hand, and also whether errors in grammar are systematic and posing a threat to communication.
Discourse capability -25%	This domain poses the question as to whether the CSR is able to explain information/instructions clearly and logically.
Interactional and strategic capability -25%	This domain poses the question as to whether the CSR is able to build a relationship and interact effectively with the interlocutor on the phone.

Table 2). This is where the language related to company policies, products and processes is used as described below:

The tasks assessed are authentic live and recorded calls carried out by QASs on the floor.

Adapting the ‘weak’ version BUPLAS recruitment assessment to a ‘strong’ version BUPLAS QA for appraisal and coaching purposes, has required close collaboration with SMEs over the years and the adaptation has involved three main considerations:

- (i) Re-weighting the value ascribed to the existing BUPLAS VA criteria towards a ‘strong’ LSP performance assessment for on-the-floor.
- (ii) Adding a 5th domain of task fulfillment, which BUPLAS CA calls ‘Business Solutioning’; this probes the handling of the business product information and handling of the processes for the specified account.
- (iii) Elaborating the performance domains to be more analytical thus providing better diagnostic information for feedback and coaching.

It is important to note that each business account may choose to ‘weight’ the criteria towards those that are of particular value to the business. For example, CSRs are the international helpdesk of a credit card company will require good interactional skills to deal with anxious customers whilst CSRs may need excellent discourse skills to explain how to install a new IT product for an IT account. It is also interesting to note that SMEs viewed the ‘lexico-grammatical’ and ‘pronunciation’ domains as less critical for the purposes of QA as they assume that recruitment assessment (LSP assessment FOR work) has eliminated the risk of there being major language proficiency problems on the floor.

If they are coming out very low in these two domains then recruitment has let them slip through the net...that’s not to say they don’t need attention, but it is much more likely that the other domains need to be coached for. (Indian female QAS manager-mid 40’s, 2013)

Table 2 BUPLAS as quality assurance

BUPLAS quality assurance criteria Call Assessment (CA)	Focus
Task fulfillment/Business solutioning - 35%	This domain poses questions around the subject matter knowledge and skills displayed by the CSR in his/her ability to resolve customer problems in a timely and professional way. Contains product and process knowledge as informed by SMEs.
Discourse capability – 25%	This domain poses the questions as to whether the CSR keeps control of the call and is able to explain business information and processes clearly and logically to the customer.
Interactional and strategic capability -25%	This domain poses the questions as to whether the CSR is able to profile the customer successfully and build a relationship and interact effectively according to the context.
Lexico -grammatical choices -10%	This domain poses the questions as to whether the choice in the grammar and vocabulary is sufficient for the communication task at hand and also whether errors in grammar are systematic and posing a threat to communication.
Pronunciation -5% (+ intonation, word stress)	This domain poses the questions as to whether the agent is globally comprehensible and is using the prosodic features to make meaning appropriately.

The ‘knowledge and skills’ transfer of BUPLAS CA onto the floor takes longer than that of recruitment. As well as the two-day BUPLAS VA assessment training workshop, QA and other SME personnel responsible for using BUPLAS CA as a formative tool (i.e. for diagnostic, feedback and coaching purposes) are also required to attend an additional five-day workshop to ensure diagnoses of communication problems are well identified and prioritised. This workshop also introduces and develops skills for communications coaching.

Weak and strong LSP performance assessment characteristics

Having described both the nature of the two business points within the call centre and the assessment tool tailored to meet their respective needs, I now synthesize the different characteristics of each under definitions ‘weak’ and ‘strong’ LSP performance assessment. The distinctive characteristics of each revolve around the purpose and impact value, the location, the task-type, the rater and the number of the criteria (including the weightings ascribed to each, and level of elaboration) as tabulated below (see Table 3).

It could be said, as suggested by McNamara 1996, that BUPLAS CA (quality assurance and coaching) reflects what he called the ‘real life’ context of the call centre in a stronger way than BUPLAS VA (recruitment) in that it uses ‘real life’ SMEs as assessors, it uses ‘real life’ tasks in the ‘real life’ context of work. Furthermore, it has an added ‘indigenous’ domain ‘Business Solutioning’ that has been collaboratively developed with the SMEs and which is likely to be more heavily weighted than the other criteria in appraisal and feedback.

BUPLAS can best be described as an ‘end-to-end’ assessment solution tailored to, and embedded into the call centre workplace. Whilst BUPLAS CA for quality assurance has an added criteria probing business product and processes information and compliance, the other 4 criteria remain the same but are differently weighted and described to better fit their distinct purposes. The advantages of having common BUPLAS criteria across the worksite has ensured a common metalanguage and a shared understanding of what good customer service communication in English may entail. Fragmented, and at times contradictory beliefs within the call centre worksite,

Table 3 ‘Weak’ and ‘strong’ LSP performance assessment for Asian call centres

LSP assessment characteristics	Weak LSP performance assessment FOR work: recruitment BUPLAS Voice Assessment (VA)	Strong LSP performance assessment AT work: quality assurance BUPLAS Call Assessment (CA)
Purpose: the usual business usage	Employment entry: recruitment and/or reassignment	Quality assurance: appraisal performance, coaching and support
Location	Recruitment	On-the-floor at work
Task type	Phone interview task to check pronunciation and level of language proficiency.	Authentic exchange-‘real work’ calls.
Rater	Non-SME assessor e.g. recruiter or communications trainer who assess making interferences and predictions for success	SME assessor e.g. quality specialist or team leader who assess by making direct observations of authentic calls.
Criteria, number of criteria and weightings	4 domains-no product/process knowledge assessed Content knowledge peripheral	5 domains-5 th domain assesses product and process knowledge Content knowledge central

and definitions about what good communication is, has been a common complaint within this industry (Lockwood 2012) over the years.

Discussion and conclusions

In this article I have described the emergence of the BUPLAS LSP spoken performance tool and set of processes that have been developed over the last decade whilst researching and working with major call centre worksites in Asia. I have argued that BUPLAS (VA) for work reflects a 'weak' version of assessment tailored to recruitment needs, whilst BUPLAS (CA) at work reflects a 'strong' version of assessment tailored to the business quality assurance needs, a key one being coaching feedback. The former is primarily summative in focus, whilst the latter is primarily formative. A number of characteristics that flow from this distinction include how the criteria are elaborated, prioritized and weighted; who does the rating, and the nature of the assessment tasks. Key to the building of this model of assessment within call centres has been the high commitment of the business to develop and use assessments that are 'fit for purpose,' however the generalizability of such a model across other worksites is unknown.

One interesting LSP assessment approach cited earlier in this article, is the work currently being carried for the OET assessment for health professionals in Australia. The OET assessment in its current form is one test (covering productive and receptive skills) and aims to meet the needs of entry recruitment by determining the candidate level of language and hopefully predicting quality success at work (i.e. in the health professional/patient interaction). Other LSP performance assessments (e.g. aviation, vets, teachers) appear to be using similar models. A question arises whether there is also a need for on-the-job assessment in these professional settings to ensure quality and on-going improvement.

A key issue for BUPLAS reliability, where knowledge and skills transfer is a non-negotiable business requirement, has been what appears to be success in training non-language background raters to assess the quality of English communication at work. Unlike most other LSP assessment tests, SMEs are used for all BUPLAS assessments. Over the last ten years, call centre recruiters and quality assurance personnel have shown acceptable rates of reliability after attending the requisite BUPLAS training workshops, and after multiple and regular calibration sessions. As part of BUPLAS accreditation, all candidate calibration results are recorded so their performance over time is tracked and they are accredited as calibrated once they are consistently within the calibration threshold (.5 measure on BUPLAS). These findings perhaps contest the common assumption that only language assessors can reliably use language communication scales and descriptors; SMEs with appropriate training and support appear to be doing an equally good job. In the recent OET studies (reported by McNamara in Cambridge ESOL presentation, 2013 www.policyreview.tv/conference/920-Cambridge-english-centenary-co) it was revealed that the new SME informed 'indigenous criteria' of 'discourse management' and 'clinician engagement' were reliably scored by language assessors after training and calibration. Further studies on how the background of the rater affects the design and reliability of the tool, and how such raters can be trained and calibrated are all areas for further and rich investigation.

A further question that has emerged in LSP 'strong' BUPLAS CA assessment processes has been the number of calls to be assessed to make a QA assessment. SMEs

propose using multiple calls for the purposes of quality assurance, diagnosis and feedback at work. They claim using a random selection of just one call as the basis for assessment is both unfair and may be unrepresentative of their real communication ability on the phones. It was therefore decided by the business that basing assessments on a minimum of three calls was desirable as this would allow patterns of problems over these calls to emerge thus providing a rich diagnosis opportunity prior to each coaching session. It would therefore seem that the number of 'tasks' assessed vary in recruitment (one phone interview) whilst on-the-floor three assessments are recommended for appraisal and coaching.

An interesting issue that has arisen in developing the strong BUPLAS CA version has been the role of the interlocutor when assessing the real calls. Where the interlocutor is the customer and where, according to the business, the customer is always right, the resulting power differential (and mood of the customer) can potentially negatively impact the CSR assessment. If we believe communication to be socially constructed and a joint responsibility, then as McNamara (1997) rightly asks 'who are we to blame if communication goes awry?' (p. 459). This aspect of the social dimension of assessment warrants further research in this context where an angry unreasonable caller may be seen to negatively impact the CSR's ability to communicate.

Given that BUPLAS CA probes both linguistic and business informed criteria, the inevitable question has already been raised as to the efficacy of using this strong form of assessment with native - speaker CSRs working in on-shore call centres in the UK and USA. This question has also been raised in the context of the aviation test entry test (Alderson, 2011). A study has recently been proposed by a large European bank to BUPLAS CA assess 400 (200 native and 200 non-native) authentic calls to establish if there is any merit in using BUPLAS CA for appraisal and coaching feedback purposes for native speaker CSRS onshore (i.e. in the UK, USA and Australia).

A final question that this study raises is whether there are parallel definitions for LSP performance assessment in writing and 'chat' support in this industry. Much of the IT support in call centres is now carried out through synchronous and asynchronous email and chat. Many customer complaints relate to poor customer relationship building using inappropriately long responses that appear to have been cut and pasted from text templates provided by the business. This would provide a further rich area for LSP writing assessment research for work and at work.

The call centre industry, both on-shore and off-shore, values English language communication as its 'core commodity' and therefore in this sense has provided an ideal site for LSP spoken performance assessment research and development. Business stakeholders in the Asian context have constantly compared the quality of its onshore and offshore call centre capability, and business development, they claim, has been jeopardized by perceptions related to English language capability in its second language speaker call centre workforce. Therefore there has been high motivation and investment on the part of the business in improving recruitment and on-the-job communications assessments. The resulting collaborative processes between applied linguists and call centre personnel has been on-going, each party taking risks in the experiments of pushing the boundaries of what is possible in LSP workplace assessment design for both now and in the future and ultimately what impact such tools and processes have on the business.

Competing interests

BUPLAS is now a commercial tool used by Future Perfect Business English Specialists, a Manila based consultancy founded by me in 2004.

Received: 24 June 2014 Accepted: 9 July 2014

Published online: 04 February 2015

References

- Alderson, C. (2011). The politics of Aviation English testing. *Language Assessment Quarterly*, 8, 386–403.
- Argote, L., & Ingram, P. (2000). Knowledge transfer: a basis for competitive advantage in firms. *Organizational Behaviour and Human Decision Processes*, 82(1), 150–169.
- Bachman, L. (1990). *Fundamental Considerations in Language Testing*. Oxford: OUP.
- Bachman, L., & Savignon, S.J. (1986). The evaluation of communicative language proficiency: a critique of the ACTFL oral interview. *The Modern Language Journal*, 70(4), 380–390.
- Canale, M. (1983). From communicative competence to communicative language pedagogy. In J Richards & R Schmidt (Eds.), *Language and Communication* (pp. 2–27). New York: Longman.
- Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics*, 1, 1–47.
- Davies, A. (2010). Language Assessment in call centres. In G Forey & J Lockwood (Eds.), *2010 Globalization, Communication and the Workplace: Talking Across the World* (Continuum, pp. 242–251).
- Douglas, D. (2000). *Assessing Languages for Specific Purposes*. UK: Cambridge University Press.
- Douglas, D. (2001). Language for specific purposes assessment criteria: where do they come from? *Language Testing*, 18(2), 171–185.
- Elder, C. (2001). Assessing the language proficiency of teachers: are there any border controls? *Language Testing*, 18(2), 149–170.
- Elder, C, Pill, J, Woodward-Kron, R, McNamara, T, Manias, E, Webb, G, & McColl, G. (2012). Health professionals' views of communication: implications for assessing performance on a health-specific english language test. *TESOL Quarterly*, 46(2), 409–419.
- Forey, G, & Lockwood, J. (2007). I'd love to put someone in jail for this' An Initial Investigation of English needs in the Business Processing Outsourcing (BPO) Industry. *English for Specific Purposes*, 26, 308–326.
- Friginal, E. (2013). Evaluation of oral performance in outsourced call centres: an exploratory case study. *English for Specific Purposes*, 32, 25–35.
- Halliday, MAK. (1985). Context of Situation. In MAK Halliday & R Hasan (Eds.), *Language, context and text: Aspects of language in a social-semiotic perspective* (pp. 3–14). Victoria: Australia:Deakin University Press.
- Hamp-Lyons, L., & Lumley, T. (2001). Assessing language for specific purposes. *Language Testing*, 18, 127–132.
- Hasan, R. (1985). The structure of a text. In MAK Halliday & R Hasan (Eds.), *Language, context and text: Aspects of language in a social semiotic perspective* (pp. 52–69). Victoria, Australia: Deakin University Press.
- Hood, S, & Forey, G. (2008). The interpersonal dynamics of call centre interactions: constructing the rise and fall of emotion. In *Discourse and Communication*.
- Hymes, DH. (1973). Towards linguistic competence. In *Texas Working Papers in Sociolinguistics* (16th ed.).
- Jacoby, S, & McNamara, T. (1999). Locating competence. *English for Specific Purposes*, 18(3), 203–241.
- Lockwood, J. (2010). What causes communication breakdown in the call centres? The discrepancies in BPO workplace training and academic research. In G Forey & J Lockwood (Eds.), *Globalization, Communication and the Workplace*. Continuum International Publishing Group Ltd.
- Lockwood, J. (2012). Are we getting the right people for the job? a study of English Language recruitment assessment practices in the Business Processing Outsourcing (BPO) sector: India and the Philippines. *J Bus Commun*, 49(2), 107–127.
- Lockwood, J, Forey, G, & Price, H. (2008). Englishes in the Philippine Business processing Outsourcing Industry: Issues, opportunities and research. In MLS Bautista & K Bolton (Eds.), *Philippine English: Linguistic and Literary Perspectives* (pp. 157–172). Hong Kong: Hong Kong University Press.
- Lockwood, J, Forey, G, & Elias, N. (2009). Call centre measurement processes in non-english speaking contexts. In D Belcher (Ed.), *English for Specific Purposes in Theory and Practice* (pp. 143–165). USA: University of Michigan Press.
- McNamara, T. (1996). *Measuring Second Language Performance*. UK: Longman.
- McNamara, T. (1997). 'Interaction' in second language performance assessment: whose performance? *Applied Linguistics*, 18(4), 446–466.
- McNamara, T. (2013). Cambridge, UK: Cambridge ESOL. www.policyreview.tv/conference/920-Cambridge-english-centenary-co
- The Business Processing Association of the Philippines (BPAP) Report (2006). www.bpap.com Retrieved July 19th 2007.
- BPAP (2006) P2O and BPAP Periodic Survey #2 *Constraints to Growth –Summary of results p.6*.
- Ventola, E. (2005). Revisiting service encounter genre: some reflections. *Folio Linguistica*, 39(1–2), 19–43.