
Are We Getting the Right People for the Job? A Study of English Language Recruitment Assessment Practices in the Business Processing Outsourcing Sector: India and the Philippines

Journal of Business Communication

49(2) 107–127

© 2012 by the Association for

Business Communication

Reprints and permission: <http://www.sagepub.com/journalsPermissions.nav>

DOI: 10.1177/0021943612436975

<http://jbc.sagepub.com>



Jane Lockwood¹

Abstract

Although the business processing outsourcing (BPO) industry in India and the Philippines is keen to employ good English language communicators, especially to work in call centers dealing with native speaker customers, it is by no means certain that the investments they are making into their language recruitment assessment processes are paying off. Many call centers are reporting recruitment conversion rates as low as 1%. It is further reported that the lack of English language skills is the main reason for rejection. But how exactly are the call centers carrying out their English language recruitment assessments? How effective are the BPO industry's recruitment assessment tools and practices? This article will report on a study of English language recruitment assessment practices carried out in two multinational companies (MNCs) that have BPO sites in India and the Philippines. One of these MNC sites offers IT and back office support and the other operates a global network of call centers. Research access was granted to their human resource departments to shadow their recruitment language assessment practices, to interview key stakeholders, and to collect and analyze relevant documentation. The study revealed a problematic stakeholder understanding of what to look for in language ability when

¹City University of Hong Kong, Kowloon Tong, Hong Kong SAR

Corresponding Author:

Jane Lockwood, City University of Hong Kong, Kowloon Tong, Hong Kong SAR

Email: lockwood@cityu.edu.hk

recruiting staff; it also revealed the problematic use of language assessment tools and practices in terms of validity, reliability, practicality, and fairness. It will be argued in this article that this unfortunate combination may be resulting in unreasonable language assessment "gatekeeping" to BPO industry employment and that the problems being experienced by the BPO industry to recruit enough good speakers of English may reside more in their own practices than in the levels of English of their graduate applicants.

Keywords

business processing outsourcing (BPO), English language assessment, business English as a lingua franca

The business processing outsourcing (BPO) industry is one of the fastest growing global industries, and over the next 15 years, it is estimated that more than 3 million U.S. service industry jobs and US\$136 billion in wages are expected to move to countries like India, Russia, China, and the Philippines (O'Neil, 2003). The global outsourcing market is worth in excess of US\$600 billion annually, including the United Kingdom and Europe, and new outsourced and offshored destinations are unfolding in the developing world where an educated workforce is readily available. For the call center sector of this industry, spoken language skills are key.

The BPO industry in India and the Philippines, particularly in this call center sector, requires excellent English language phone skills when working with native speaker customers. However, this industry has been consistently reporting difficulties in hiring good communicators due to the lack of English language proficiency skills (Business Processing Association of the Philippines [BPAP], 2006). In a recent study in the Philippines (BPAP, 2006), 89% of industry respondents reported that English had a significant impact on their organization's ability to grow. The report makes the concluding remark that

the results of the survey show that English proficiency is a key impediment to growth. (BPAP, 2006, p. 11)

Having already carried out a number of English language communications consultancy assignments in the BPO industry in India and the Philippines over the past 7 years, it seemed to me that the low conversion rate at recruitment had more to do with the kinds of English language assessment tools and processes being used at recruitment than the poor level of English of the applicants. I decided therefore to embark on a research study to investigate what key stakeholders perceived "good communication" to be as part of an investigation of the tools and processes they use at recruitment. The value of such a study, it was hoped, would alert multinational companies to the

requirement for valid, reliable, practical, and fair recruitment assessment practices, which may, in the long run, mean the conversion rates would improve.

The BPO industry in India and the Philippines, particularly in this call center sector, requires excellent English language phone skills when working with native speaker customers.

The Research Questions and the Literature

Understanding what key stakeholders define as “good proficiency” and how they assess it were the two key research questions in my study. The specific research questions were as the following:

Research Question 1: What do BPO key stakeholders understand good communications assessment to be?

Research Question 2: What tools and practices are being used?

Two BPO worksites were selected for investigation. These worksites represent two different examples of BPO work: one a back office and IT support center and the other a call center. These are referred to as Company A and Company B, respectively, for confidentiality reasons. Company A is domiciled in the United Kingdom and Europe and has back offices in the Philippines as well as in India. This company provides IT and back office support for a range of internal and external clients across the world, with English being the dominant corporate language. Although this company runs operations in both India and the Philippines, only the Philippine office was involved in this study. English is the business lingua franca in this company, and employees are often required to communicate in English with their counterparts in the United Kingdom, Europe, Asia, and Australia, where English may not be the first language. The Business English as a Lingua Franca (BELF; Louhiala-Salminen, Charles, & Kankaanranta, 2005) context in Company A is acknowledged but not fully addressed in this article, as English language assessment research is scarce in this regard. Company B is a large financial institution domiciled in the United Kingdom. This organization has more than 100 customer service transaction accounts supported by call centers across South and East Asia. The study in Company B was carried out in India and Manila.

As a BPO industry group, call centers and back offices can range significantly in their communications requirements. Where call center agents simply have to look up directory enquiries and provide numbers to clients, the communication needs are very different in a financial advisory account where the agent is communicating with clients who may be investing significant amounts of money and requiring advice.

In Company B, there was a range of communication requirements commensurate to the complexity and sensitivity of the different customer accounts. In Company A, IT

support staff are in regular communication with overseas clients and colleagues. The communications assessment practices and tools therefore need to reflect these complexities and differences and also need to have processes that “feed in” employees at the right level of English for the very different post requirements and audiences with which they will communicate.

The English language assessment literature is rich in discussions about the relative merits of using different applied linguistic frameworks to achieve construct validity and reliability in the design and use of any test (Alderson, Clapham, & Wall, 1995; Bachman, 1990; Luoma, 2004; Weir, 1998). Much of the discussion has revolved around pushing criteria for assessment beyond the traditional preoccupations with pronunciation and grammatical accuracy to a consideration of broader communicative domains such as discourse and interactive capabilities (Bachman, 1990; Bachman & Palmer, 1982, 1996; Bachman & Savignon, 1986; Canale & Swain, 1980; Davies, 1988; Hughes, 1989; Weir, 1998). English language assessment practices that reflect communicative approaches to language training in Language for Specific Purposes (LSP) contexts are also well covered (Douglas, 2000, 2001; Elder, 2001; Lumley & Brown, 1996; McNamara, 1996, 1997), and well-documented accounts of test development for different groups of occupations and professionals abound. McNamara (1996) and McDowell (1995) have both developed standardized and competency-based tests for teachers and health professionals in Australia; Douglas (2000) looks at specific target language use situations to develop test content and test methods for highly specific LSP, such as English for air traffic control. Such frameworks, however, have yet to be fully understood in an industry context and incorporated into business language assessment practices on-site.

The discussions to date are confined mostly to English language assessment researchers working in either generic applied linguistic contexts (e.g., schools, colleges, and universities) or in relation to specific occupations (e.g., teachers, doctors). In the latter case, the testing requirement has been largely to provide an LSP test for benchmark entry purposes into the profession. So how is the BPO industry different? There appear to be two main differences. First, as described earlier, the entry requirement into BPO worksites is dictated by the business communication requirements set by the different accounts/types of work, so there will necessarily be *multiple entry benchmark requirements* for the industry depending on the type of work and the range of accounts. Second, the BPO industry, and in particular call centers, regard good communication in English as a *key core skill* for working on the phones. The call centers therefore require *ongoing evidence* of communications improvement and competence, requiring an *end-to-end English language assessment solution*. The quality assurance processes in the call centers are driven by metrics and agents are continuously and quantitatively tracked for English language communication quality (Forey, 2010; Frigal, 2007). This means that there are complex requirements for language assessments that are linked to specific work and accounts, and there are added requirements for ongoing measurement for quality assurance and training/coaching support during the employment span of the employee within the BPO worksites (Lockwood, 2008). These business needs result in

key stakeholders in the industry (HRO personnel, trainers, team leads, account managers, and quality assurance personnel) requiring good levels of English proficiency themselves; a shared understanding of what good English language skills comprise and alignment on the how to measure them. Most language assessment studies to date, even in the area of LSP assessment, have focused on the test product rather than on the business processes and contexts in which they occur (Lockwood, 2010). In the BPO industry, the business requires a multidimensional and an ongoing English language assessment framework that constantly informs management about the development and quality of the communication effectiveness of its employees as English communication, particularly in call centers, is a highly valued key core skill.

Much of the discussion has revolved around pushing criteria for assessment beyond the traditional preoccupations with pronunciation and grammatical accuracy to a consideration of broader communicative domains.

Although the applied linguistic literature is rich on English for Specific Purposes training needs analysis, program development, and genre analysis (Bargiela-Chiappini, Nickerson, & Planken, 2007; Bhatia, 2008; Dudley Evans & St. John, 1998; Flowerdew & Wan, 2010), the literature is scarce where specific language assessment needs analysis is concerned, particularly in the area of business worksites. Reeves and Wright (1996) address the specific issue of front-end language requirements of workplace posts through what they have coined “a language audit.” They take a broad-angled ethnographic view of needs analysis in the workplace and encourage language providers to “tailor” courses specifically to the needs of the workplace based on language benchmarking the “work” within the organizations. Once these language needs and thresholds are understood, planning for language improvement and measurement can take place.

A language audit therefore should strive to help a company’s management make the right strategic decision in recruitment, in modifying the organization and the behaviour of some departments, as well as allocating resources for training and quality assurance. (Reeves & Wright, 1996, p. 2)

Revell (1994) talks specifically about language testing in the Hong Kong corporate context and makes an interesting and relevant link between language assessment and quality management practices in Hong Kong companies, where he says

internal and external customer requirements are first priority, but other sources of information may be needed and should be found to make sure training and assessment are aligned with the company’s quality requirements. These may be found in training criteria, in job descriptions, or in external quality requirements. (p. 313)

However, although Revell (1994) has articulated the problem, he does not offer any solutions in the form of practical guidance or suggests any prevailing theoretical frameworks that may prove useful to the English language workplace assessment

practitioner. Lockwood (2002), in a further analysis of workplace training and evaluation processes in Hong Kong workplaces, has argued for an interdisciplinary approach to curriculum and program evaluation incorporating applied linguistic and business management models (Bramley & Pahl, 1996; Goldstein, 1993; Kirkpatrick, 1994) to ensure business stakeholder involvement in these educational processes. This approach relies on educationalists taking a "consultative role" and an ethnographic research type in the workplace by gathering key stakeholder views and understanding the business requirements and culture in order to ensure that training and assessment have an appropriate impact on workplace performance. These interdisciplinary and ethnographic research approaches have been key to understanding the business requirements for language assessment.

The Study

An ethnographic study in both sites yielded qualitative data that form the basis of this study. Having consulted in many similar BPO recruitment sites in the past in Manila, I felt it was important for the purposes of this study to replicate the kind of study I had carried out in previous consultancies and also to replicate the kind of data I had been able to gather previously. This would involve spending time at each site observing their recruitment practices and making notes and recordings where relevant and interviewing key stakeholders and gathering relevant documentation for later analysis (e.g., language recruitment assessments used, recruitment forms, statistics on conversion rates). Gathering these data involved six visits to the BPO sites. Two visits were made to Company A (nonvoice BPO), two visits to Company B (voice BPO) in Manila, and two visits to Company B in India. The first visits comprised interviewing the key recruitment stakeholders at each site, namely, three HRO managers, five trainers, seven recruiters, three quality assurance personnel, and four supervisors. These were carried out individually and through small group focus discussion groups. In Company A, I was also able to interview the chief executive officer (CEO) to ascertain his views on the communication needs of his employees in Manila and his perceptions on how best to screen for good English language skills. Each interview/focus group discussion was conducted for about 45 minutes and recorded. Although I had devised a set of questions, the interviews/focus group discussions were semistructured, where interviewees and group participants were invited to expand on areas of importance and relevance to them as they emerged in the discussion. These were then transcribed and analyzed for patterns of concern around the research questions earlier outlined. It was felt to be important to talk to as many as the key stakeholders as possible in the BPO sites to triangulate the findings and to note differences in perceptions about the importance of these language recruitment assessment tools and processes within the organization. Unfortunately, I was not able to interview any of the applicants themselves. This perspective would have provided another important dimension in this kind of study.

The purpose of the second visits was to “shadow” the recruitment processes at face-to-face and phone interviewing sessions. In Company B, I was able to observe both the initial interviews conducted by the HRO department recruiters (six interviews) as well as a small number of the final interviews conducted by the supervisors (three interviews). I was also permitted to complete some of the assessments myself in order to experience and evaluate the assessment screening tool and procedure used by Company B. This was made possible by providing me with a candidate password. Shadowing the phone interviews in Company A was made possible by inserting a “jack” into the phone for the researcher to listen in. Relevant documents were also collected during these visits, namely, copies of tests where they were available as well as the hiring, scoring, and reporting documentation. I was also able to view the computerized recruitment system for communications assessment that converted all the interview data into a set of numerical profiles for each recruitee in Company B.

These interdisciplinary and ethnographic research approaches have been key to understanding the business requirements for language assessment.

Findings

The findings below are based on the analysis of the data gathered across both companies and across the three sites. For Companies A and B, the findings are divided into three themes that emerged from analyses of the data as they relate to the research questions. First, “The tools and processes used for making language assessment judgments,” second, “Problems in awarding scores,” and third, “An overall evaluation of what we do.” The final discussion will make summary comments on patterns and concerns expressed in the findings of these studies, the limitations of the study, and suggestions for further research.

The Nonvoice Site in Manila: Company A

The tools and processes used for making language judgments. Like many BPO work-sites, the recruiters at the nonvoice site of Company A in Manila reported that they have a dual purpose for the first job interview. The interview is normally done over the phone after the company has received the job application and lasts for about 15 minutes.

The process we have is really that our recruitment staff have an initial interview and this short interview would include the assessment of communication skills and the behavior interview, so it’s sort of you are looking at the suitability as well as you know how they could communicate with us. (Recruitment officer in Company A)

The “rapport-building” questions that they ask at every interview are the following:

1. Tell me about yourself.
2. How do you spend your free time?
3. How do you see yourself 5 years from now?
4. Under what conditions and in what environment do you work best?
5. Why do you want to work for Company A? What do you know about our company? (Example Test 1)

These questions are expected to generate extensive answers lasting for about a minute.

When asked what criteria they use to judge the level of communication, the recruiters reported that the most important one is “grammar.” When asked what they meant by “grammar,” one of the recruiters replied,

Grammar and tenses like you know when referring to past incidents, past tense in that aspect. If their sentence formation are in the right context, that will also include, you know “quick thinking,” meaning that if we ask them a question they are able to formulate their thoughts immediately, we look into both quickness and accuracy. (Recruiter in Company A)

The recruiters felt that *quickness of thought and response* was a valued communication feature in the work at Company A, where middle and senior IT specialists need to talk to their counterparts in the United Kingdom and Europe in real-time phone and video conversations.

It can be really embarrassing and frustrating when our IT people go into conference calls with U.K. and Europe, they completely close down and their silence and long hesitations are wrongly construed as they don’t know, they appear unprofessional, so it’s important to make sure they can stand up to this and show that they can respond and participate. (Middle manager in Company A)

When probed further, the recruiters also talked about their understanding of coherence:

We are looking for the coherence in terms of their thoughts. They could be speaking about a topic in the first two sentences and then jumping onto a different explanation. We look at lapses and hesitations, you know when they are formulating their thoughts, we mark down if they hesitate. (Recruiter in Company A)

On the topic of pronunciation, the recruiters appeared to have a low tolerance for first language interference, and mother tongue “accent” seemed to greatly influence decisions for hire.

Especially here in the Philippines, people coming from different regions and provinces, they have different accents, they pronounce the words wrongly so we pay attention to that as well. This is important because interaction with our clients and counterparts in other countries is a big part of their job so that it is very important that they could express themselves and they could be understood. Because most of the discussions are done over the phone and there are conferences which are on the speaker phone, so it is difficult if you can't speak your words properly. (Recruiter in Company A)

This emphasis on the mother tongue interference (MTI) was refuted by the CEO and one of the other native speaker expatriate managers. The CEO said on this point,

We are all “over” [finished] having our voice and nonvoice BPO employees sounding as if they are native speakers, thankfully the industry has matured on this front and we realize that there will be MTI in how our employees speak and that's fine as long as they are comprehensible. After all they are talking with nonnative speakers in Europe all the time, French, Dutch, Scandinavians, so who is having difficulty understanding who and for what reason? Generally educated graduates in the Philippines are pretty easy to understand. (CEO of Company A)

Problems in awarding scores. When asked how the applicant to Company A is formally scored for language ability there appeared to be no formal criteria, and one of the recruiters said,

Most of the recruiters are in the same level in terms of their own language and how they speak, we are good communicators ourselves so we know what is a good communicator, we all agree and if we agree it is OK. Usually the problems are too many lapses and poor grammar and how they sound, meaning you rejected them because they made too many mistakes in grammar or pronunciation. (Recruiter in Company A)

The final interview form at Company A has a “communication” component that lists the key indicators as follows:

1. Exhibits good diction and voice quality
2. Uses appropriate grammar and vocabulary
3. Expresses ideas clearly and concisely
4. Listens attentively to interviewers responses (Example Test 2)

This list however did not have a rubric to language assess against.

An overall evaluation of what we do. When asked if they were satisfied in Company A with the speaking assessment process, they offered a number of suggestions for

improvement. First, they felt that the questions being asked were both “tired” and posed a security risk as applicants circulate what will be asked in the interview among their friends. The recruiters try to overcome this by asking different questions, but they reported difficulty in being able to do this well. The HRO manager said,

We want to have the recruiters go for regular training for interviewing skills, it’s not just what questions and how to ask them, it is the way you formulate the questions, you know not just getting yes or no answers from the candidates. (HRO manager in Company A)

Other recruiters suggested having a benchmark level specified as is the case for their IT specialists. Connected to this suggestion, they could see that they would need sets of criteria and descriptors:

Actually we have a lot of flaws in our current practices, rather than just writing a comment, we should already have a benchmark or a level of specific description. At present we just write a comment. (Recruiter in Company A)

There were some practical concerns also articulated in the interviews at Company A about the initial phone interview where there was no prior warning to the applicant that this speaking assessment would take place. In shadowing the process, although there were six candidates to talk to that afternoon, the interviewer was only able to locate two of them. One of these candidates was in a shopping mall having lunch and the noise and timing of this high-stakes speaking assessment called into question the validity, the reliability, and the fairness of such practices. The quality of the sound in terms of background noise was very problematic in my view in determining the true communication level of the candidate. It seemed odd, if not unfair, to conduct such a high-stakes speaking assessment at random and under such potentially difficult circumstances.

Company A also conducts a writing test at recruitment as many of its employees need to write emails to business counterparts overseas. The HRO recruiters felt that they did not have the internal expertise to make these judgments about writing competency and therefore an external provider was sourced to provide this service. The tool is an online writing assessment that takes about 45 minutes to complete and costs US\$35.00 per person. It is a multiple choice test and does not, in fact, require the candidate to produce any writing at all. The test comprises reading comprehension passages followed by multiple choice questions. The other exercise relies on the test taker being able to spot the grammar mistake. For example,

Which of the following sentences contains a grammatical error?

- a. A ten-mile walk is no joke in the rain
- b. A man’s elephant walk down the street
- c. It is good to wear a happy smile

- d. An healthy appetite is good for you
- e. The lions search high and low for their prey. (Example Test 3)

The recruiters were uncomfortable with this type of test; first because the items were at times ambiguous (in the above example two were grammatically incorrect in my view) and second because it did not actually test writing. The recruiters said candidates consistently performed poorly on this test and it provided no diagnostic information for the workplace. When the native-speaker CEO (with an MBA and an LLB from Australia) was alerted to the questionable validity of such a test for writing, he agreed to complete the test himself. He was benchmarked “an intermediate level writer requiring support.” Needless to say, the test was pulled immediately.

It seemed odd, if not unfair, to conduct such a high-stakes speaking assessment at random and under such potentially difficult circumstances.

The Voice Worksite: Company B

The tools and processes used for making language judgments. In contrast to the example above, the second workplace (Company B) had developed benchmark levels, encompassing “high,” “medium,” and “low” voice for its call center work. They had put much time and many resources into devising an array of different site-based speaking assessment solutions. This in itself, however, was viewed as problematic by the business as the site-based initiatives were not regionally shared and each had developed its own solution to the problem of measuring language proficiency at recruitment and for communication success in QA. For Company B, I will report on the speaking assessment tools and processes in both the Philippines and India where I was able to collect data, interview recruiters, and shadow the recruitment processes. These sites operate call centers for a large financial services organization supporting U.K. and U.S. customers and clients. Interestingly, despite the fact that these worksites are part of the same company, they carry out different processes and use different tools in each site. Company B saw this ad hoc emergence of different tools and processes as problematic in terms of continuity across the region and in terms of resources (everyone in each site was reinventing the wheel). From a language assessment point of view such diverse practices would also appear to be problematic in terms of ensuring validity, reliability, and practicality where there needed to be an “across the region” understanding and meta-language for talking about English language communication levels.

Both sites have “walk in” spaces in their recruitment areas where candidates will literally walk in off the street for a communications interview as the first part of the recruitment process. The candidate is first asked to complete a U.K. developed computerized placement test, called their “Quick Placement test.” To move onto the next stage, she/he must get a threshold score. It was noted that some potentially good “speaking” candidates were screened out at this stage. I completed the test myself and

found the test to be very British in orientation. It probed language choices that may be unfamiliar to Filipinos; for example, from a series of pictures, the candidate had to select “the girl wearing a frock.” “Frock” is a British word (although even now old fashioned in the United Kingdom) and not easily understood by young Filipinos who watch mostly U.S. television and movies. Part of this test involved a Cloze test with high ambiguity in the selection of the correct vocabulary items as shown below:

Dull/plain gloves are worn for everyday/day-to-day wear. (Test Example 4)

After completing this test successfully, the candidate was then invited to do a home-grown listening comprehension test of authentic Company B calls. This was problematic because the calls were technically very complex and the listening comprehension questions were “content” driven and devised by quality assurance specialists with no expertise in language assessment. For candidates who have no knowledge of the organization’s financial services and products, this was difficult to do, even for a native speaker. Furthermore, the questions and multiple choice items were not grammatically correct and relied heavily on knowledge of company policy.

Question 4: Which of the following would have best increase the possibility of getting the sale from the customer?

- a. The agent could have focused on the benefits that the customer could get from the product
- b. The agent could have given more effort on selling the product
- c. The agent could have been more enthusiastic in offering the product to the customer (Test Example 5)

It was notable that two native speakers from Britain who took the test at the time of this research were only able to score 67% on this test. India did not have this listening test requirement at “walk in.”

Problems in awarding scores. The candidates in both locations then moved onto the first speaking assessment, which was carried out by recruiters who are, in some cases, also the trainers. In India, a modified form of the Common European Framework (CEFR) document is used. This was prepared by an outside consultant some years ago. The Philippines site however had developed its own speaking assessment rubric. This had been constructed by trainers and recruiters on site. Both processes do not appear to have been sufficiently informed by “best practice” in language assessment insofar as the adapted CEFR, although having some appropriate domain descriptions appeared not to have been adequately understood, trained for, calibrated, nor supported within Company B. The Indian trainers complained that calibration sessions using the CEFR had become little more than an error counting exercise and consequently yielded narrow and negative data on what the candidate could not do. This was further exacerbated in the second and final interview carried out by the supervisors in India, who similarly made English language judgments by counting grammar and pronunciation

mistakes as well as counting the number of hesitations and pauses (irrespective of whether these interfered with communication). The supervisor in the company always makes the final decision on the English language communication level of the candidate, and she/he does this according to the number of grammatical and pronunciation mistakes that she/he believes to be an accurate predictor of success on the phones. The other domains in the CEFR tool, such as “fluency,” did not appear to be well understood or activated in the recruitment assessments. Although observing the recruitment interviews in India, little awareness seemed to be evident on the part of the recruiters as well as the supervisors about the power relationships inherent in the interview context, or the effect of “nerves” when collecting data to make high-stakes assessments. One trainer commented,

I know we’re not always seeing the best when they are under so much pressure, I mean many of these guys are here because this job can get their families out of poverty . . . get an education for the kids. . . . I’ve seen some applicants break down when they get to the final interview with the supervisor. (Recruiter in Company B)

The Philippine spoken assessment solution appears to be equally problematic because the rubric has poorly defined criteria with problematic overlap. This is not surprising as the rubric was constructed by trainers and recruiters with no formal Teaching English to Speakers of Other Languages (TESOL)/applied linguistic qualifications. This is common in the BPO industry where stakeholders take a common sense and intuitive approach to language assessment and write long lists of assessment criteria; the more the better it seems. The communications team in one of the Company B sites had devised a list of 12 communication “attributes” or criteria to be individually assessed as “yes,” “no,” or “not applicable” in their rubric as follows:

1. *Neutral inflection*: Description—Participant’s native inflection must not be evident.
2. *Clear and distinctive speech*: Description—Participants must not stutter, slur, or poorly string words together.
3. *Vowel and consonant pronunciation*: Description—Participant must be able to pronounce vowel sounds correctly.
4. *Professional tone*: Description—Participants must not sound monotonous but must be aware of the peaks and valleys of intonation.
5. *Moderate pace*: Description—Participant’s rate of speech must match that of the listener’s. She/he must not stumble on words or display any sense of urgency in his/her speech.
6. *Sentence construction*: Description—Participants must not use double past tenses and verb contractions must also be avoided. Parallelism must be observed. Participants must use correct verb/subject agreement at all times.
7. *Spontaneity*: Description—Participant must respond immediately to the customer query and not hesitate.

8. *Cultural phrasings*: Description—Participant should not communicate with any awkward cultural interpretations.
9. *Response cohesion*: Description—Participant must respond in sentences with complete thought. Sentences should be linguistically and pragmatically correct. Word sequencing should likewise be observed.
10. *Phone etiquette*: Description—Participant must not interrupt or talk over the customer.
11. *Active listening and paraphrasing*: Description—Verbal nods and paraphrasing must be utilized to validate understanding of the customer concern.
12. *Customer sensitivity*: Description—Participant must positively set the customer's expectations through offering an assurance statement. (Example Test 6: Company B Communications Quality Assurance Scorecard)

Not only does such a list reflect a poor understanding of the nature of spoken communication, which research shows is highly ungrammatical, full of hesitations, and repetitions, but it also has overlapping domains that makes it impossible to mark. When shadowing a recruitment interview, a native speaker was denied a call center job because she “ummed” and “uhhed” too many times as she was thinking (as many of us do) before she spoke.

In the administration of the assessments, there is also a common preoccupation in recruitment with “counting the number of errors” as a way of measuring the level. One recruiter complained,

We always start arguing about what the criteria really mean in the calibration meetings, there is so much overlap and we often cannot agree and then we end up counting the errors and arguing about what the number of errors really mean in terms of successful communication. We always go round in circles.

It's really frustrating that we have resorted to counting errors with the CEF tool, I mean as a trainer, I really do know what this framework is supposed to do, but this is what the account managers now do and they have the final veto, so we all do the same, otherwise we get into trouble and told we're not doing our job properly. (Trainer in Company B)

Company B, at the time of this research, was also piloting a language assessment solution developed by an Indian assessment company. For confidentiality reasons, I will call this test “SpeakWell.” This company currently enjoys commercial success in a number of BPO worksites. Essentially this test is a “screen” test comprising two parts—a reading passage and an impromptu speech conducted over a 5-minute duration on the phone. The description of this test says the following:

The test taker first reads out a passage provided to him online with trap words that identify Mother tongue influence on accent and then speaks for two minutes on a topic provided to him by the evaluator. (SpeakWell Test publicity)

This test claims to evaluate a test taker's spoken English skills across four domains: voice clarity, accent neutrality, fluency, and grammar. The first two parameters appear to *overlap* in focus as they are both assessing pronunciation in the broadest sense. It is also interesting that there is no attempt to probe interactive communicative competence—a core competency in call centers.

Apart from the overlap in the domains, there appears to be a preoccupation in the tool on the *counting of errors* as the sole measure of proficiency. For “accent neutrality” the description is as follows:

The passage contains a list of 44 trap words which represent the phoneme sounds in the English language. The errors depend on the speaker's mother-tongue. To rate a test taker on Accent, the number of errors the candidate makes while reading the passage (the passage contains the MTI trap words) and during the speech are counted and based on a scale, the ratings are given. A further enhancement of Version 3 of the test is the addition of fatal and non-fatal errors which specifically identifies phoneme sounds which are considered trainable. (Example Test 7)

The “fluency” domain is similarly driven by counting the errors made by the candidate.

The test taker is rated for fluency, based on *the number* and kind of pauses, fillers, repetition of words, use of the vernacular, range and appropriateness of vocabulary and sentence construction. (SpeakWell Test publicity)

Unsurprisingly, the final domain “grammar” is also driven by counting the number of mistakes made in the speaking test.

This parameter evaluates the grammatical correctness of the test taker's speech and focuses on the correct usage of verb forms (tenses and SVA), prepositions, articles, plurals, and other grammatical items. The *number* of grammatical errors in the speech is *counted* and the rating given. Any other kind/category of grammatical error is to be noted down or counted separately to give the final rating for grammar. (SpeakWell Test publicity)

Candidates are finally given a score on the mistakes they make against a level measure. What is noticeable is that to get a “high voice” score one must not have evidence of any errors. To be denied employment opportunity you need only display 8 errors. In the discussion section below I discuss why this as well as the other assessment approaches described in this section are so problematic, and why they are failing to recruit good communicators for the industry.

An overall evaluation of what we do at Company B. Recently Company B decided to revisit its communications training and assessment processes with a view to embedding one communications and assessment system across its sites in South and East Asia to ensure improved validity, reliability, and practicality.

In Company B, the speaking assessment process can take up to 4 hours depending on the number of assessments deemed necessary. The whole language assessment process was under critical review at the time of this research and this Company was looking for an outside provider to advise them on how to improve their tools and processes with a view to improving their very low conversion rates at recruitment.

Discussion

Across both BPO sites, there appeared to be a poor understanding among key stakeholders of what comprised good spoken language assessment practices, for call centers in particular (see e.g., Davies, 2010; Luoma, 2004). This is not surprising as few trainers and recruiters have undergone any kind of TESOL training or done undergraduate degrees in applied linguistics; and even if they had, English language assessment for specific purposes is more likely to have been a theoretical module at a master's level. Basic TESOL training is highly recommended for those engaged in communications and language training as it provides a fundamental understanding of the nature of language and language learning. For example, a TESOL-trained teacher is likely to be conversant with the differences in spoken and written texts. Two ways of dealing with this lack of expertise in the companies became evident. First, the sites had either developed their own solutions based on intuition, common sense, the Internet, and past experience at other BPO sites as reported in the findings; or second, call centers and BPO sites, in general, became particularly vulnerable to commercially produced assessment solutions that promised quick and tailored solutions (as was evidenced in the Company A writing test and the Company B speaking test).

In the speaking assessments used over both sites, a common set of problems emerged from the data.

Problem Practice 1: Keeping Good Communicators Out of BPO Employment Because of Poorly Understood Speaking "problems"

Research shows that normal spoken language is full of grammatical error, it is full of hesitation and fillers, and it is full of first language interference evident in the phonology, in the textual organization, and in the grammar (Luoma, 2004; McCarthy & Carter, 1997; Schegloff, 1982). So what matters? It is highly desirable that BPO language assessors, whether they are working in recruitment, training, and coaching or in quality assurance, become familiar with the research into the nature of spoken language. It would appear that the very "stuff" of spoken language that makes it adhere and contribute to good communication, such as hesitation, rephrasing, and fillers, is being negatively marked in the BPOs worksites. Ironically, the costly and lengthy language assessment processes, evident in the two sites, are very likely excluding potential employees who are excellent communicators.

Problem Practice 2: More Concern About MTI

Another worrying practice that emerged from this study is the negatively and indiscriminately marking down on MTI mistakes in grammar and pronunciation. Research into the nature of language mistakes (Lockwood, Forey, & Price, 2008) in the call centers is not attributable, in the main, to first language interference mistakes. This practice is not only feeding the unfounded prejudice of the BPO industry that MTI is causing communication breakdown, but it is keeping good communicators out of the call centers. During the study, trainers from the voice site (Company B) admitted, that in their view, MTI was not a major indicator of communication ability.

We think we have problems related to our mother tongue, like all Indians have problems with prepositions, but actually when you look at it they have no more problems with prepositions than any other second English language learner, it's just that prepositions occur frequently in your idiomatic language such as phrasal verbs etc. So that's the difficulty, idioms are difficult for anyone, not just Indians. (Trainer in Company B)

The recent literature on BELF (Charles, 2007; Jenkins, 2005; Louhiala-Salminen et al., 2005; Nickerson, 2010) or the literature on discourse and identity (Pal & Buzzanell, 2008) has not been fully explored in relation to this assessment study as the impact of the emerging literature on world Englishes is, as yet, rather underresearched as far as English language assessment practices are concerned.

Problem Practice 3: The Counting of Errors

There appears to be an over-preoccupation within the grammar and pronunciation domains of counting the numbers of mistakes. This is based perhaps on an intuition in the minds of the lay person that the number of errors equates with a level of spoken language ability. Unfortunately, this practice will mean that many candidates with good levels of spoken English may well fail the test. Unfortunately, this intuition is not based on any theory of language or language acquisition research. Conversational analysis shows high levels of grammar inaccuracy, many hesitations, and many fillers in successful spoken interaction (McCarthy & Carter, 1997; Schegloff, 1982) in conversational English. Spoken English is very different from written language and often raters need to be reminded that focusing on grammar and focusing on error alone will not yield the best speaking assessment information on their candidates (Luoma, 2004). Additionally, second language acquisitions studies point to a hierarchy of mistakes in the interlanguage of spoken English (Pienemann & Johnston, 1987), where some systematic errors persist at even a high proficiency level. Therefore, we need to know more about this hierarchy of error in the interlanguage as perhaps an indicator of language proficiency level. In summary, this strategy of counting errors for

language assessment is fraught with problems and will not reliably get good speakers into the BPO industry, and may well, as suggested earlier, serve to keep good communicators out of the call centers.

The study, although small scale, has yielded some interesting data on practices and procedures in the BPO industry, and further studies will be required to be able to “generalize” these practices. Having said that, I spent many years consulting in the BPO industry in the Philippines and India, and this formal study confirmed much of what I had previously seen with regard to English language assessment practices.

Conclusion

The validity, the reliability, and even the practicality of the language assessments being carried out in the two worksites that were part of this study are under serious question given the discussion above. For worksites to be investing so much time and effort into recruitment language assessments that appear to be theoretically and pedagogically flawed is clearly a waste of money. How do the worksites know the recruits who they are taking in are the best available? How many good English language communicators are being excluded from an opportunity to work in the BPO sector because of these problematic language assessment processes?

One of the ways that worksites may improve language assessment quality is to ensure that those involved in the assessment process have valid tools and tasks to work with, that they are trained and calibrated to ensure reliability, and that those who are developing the workplace assessments, whether they be internal employees or outsourced providers, are TESOL trained and have a good knowledge of the appropriate language frameworks for assessment development.

Another area for further investigation that would ultimately benefit the BPO industry in language assessment and training practices revolve around the popular but misguided view that MTI is the major cause of communication breakdown. This view has negatively affected both assessment and training practices in call centers where “accent neutralization” and grammar classes driven by contrastive differences have become the accepted norm. As suggested earlier in this article, taking emerging frameworks from BELF and applying them to assessment practices would, I suspect, reveal further and different flaws in current BPO language assessment systems.

With improved language assessment processes in the BPO industry, we may yet be pleasantly surprised in the future by increased conversion rates at recruitment and perhaps also by improved quality on the floor.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

References

- Alderson, J. C., Clapham, C., & Wall, D. (1995). *Language test construction and evaluation*. Cambridge, England: Cambridge University Press.
- Bachman, L. (1990). *Fundamental considerations in language testing*. Oxford, England: Oxford University Press.
- Bachman, L., & Palmer, A. (1982). The construct validation of some components of communicative proficiency. *TESOL Quarterly*, 16, 449-465.
- Bachman, L., & Palmer, A. (1996). *Language testing in practice*. Oxford, England: Oxford University Press.
- Bachman, L., & Savignon, S. J. (1986). The evaluation of communicative language proficiency: A critique of the ACTFL oral interview. *Modern Language Journal*, 70, 380-390.
- Bargiela-Chiappini, F., Nickerson, C., & Planken, B. (2007) *Business discourse*. New York, NY: Palgrave MacMillan.
- Bhatia, V. (2008). Genre analysis. ESP and professional practice. *ESP Journal*, 27, 161-174.
- Bramley, P., & Pahl, J. (1996). *The evaluation of training in the social services*. London, England: National Institute for Social Work.
- The Business Processing Association of the Philippines. (2006). *P2O and BPAP Periodic Survey #2: Constraints to Growth—Summary of results*. Retrieved from <http://bpap.org/bpap/publications/Jun3606%20O2P%20&%20BPAP%20Periodic%20Survey%202.pdf>
- Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics*, 1, 1-47.
- Charles, M. (2007). Language matters in global communication. *Journal of Business Communication*, 44, 260-282.
- Davies, A. (1988). Communicative language testing. In A. Hughes (Ed.), *Testing English for university study* (ELT Document 127). Oxford, England: Modern English Press.
- Davies, A. (2010). Language assessment in call centres: The case of the customer service representative. In G. Forey & J. Lockwood (Eds.), *Globalization, communication and the workplace* (pp. 242-250). New York, NY: Continuum International.
- 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, 171-185.
- Dudley-Evans, T., & St. John M. (1998) *Developments in English for specific purposes: A multi-disciplinary approach*. Cambridge, England: Cambridge University Press.
- Elder, C. (2001). Assessing the language proficiency of teachers: Are there any border controls? *Language Testing*, 18, 149-170.

- Flowerdew, J., & Wan, A. (2010). The linguistic and the contextual in applied genre analysis: The case of the company audit report. *ESP Journal*, 29, 78-93.
- Forey, G. (2010). Language globalization and the workplace: Education and social implications. In G. Forey & J. Lockwood (Eds.), *Globalization, communication and the workplace* (pp. 251-270). New York, NY: Continuum International.
- Friginal, E. (2007). Outsourced call centers and English in the Philippines. *World Englishes*, 26, 331-345.
- Goldstein, I. (1993). *Training in organizations: Needs assessment, development and evaluation*. Pacific Grove, CA: Brooks/Cole.
- Hughes, A. (1989). *Testing for language teachers*. Cambridge, England: Cambridge University Press.
- Jenkins, J. (2005). *World Englishes*. London, England: Routledge.
- Kirkpatrick, A. (1994). *Evaluating training programs—The four levels*. San Francisco, CA: Berrett-Koehler.
- Lockwood, J. (2002). *Language training design and evaluation processes in Hong Kong workplaces* (Unpublished doctoral dissertation). University of Hong Kong.
- Lockwood, J. (2008). What does the business processing outsourcing (BPO) industry want from English language assessment? *Prospect*, 23(2), 67-78.
- 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. 221-241). New York, NY: Continuum International.
- Lockwood, J., Forey, G., & Price, H. (2008). Englishes in the Philippine business processing outsourcing industry: Issues, opportunities and research. In M. L. S. Bautista & K. Bolton (Eds.), *Philippines English: Linguistic and literary perspectives* (pp. 157-172). Hong Kong: Hong Kong University Press.
- Louhiala-Salminen, L., Charles, M., & Kankaanranta, A. (2005). English as a lingua franca in Nordic corporate mergers: Two case companies. *English for Specific Purposes*, 24, 401-421.
- Lumley, T., & Brown, A. (1996). Specific purpose language performance tests: Task and interaction. *Australian Review of Applied Linguistics*, 13, 105-136.
- Luoma, S. (2004). *Assessing speaking*. Cambridge, England: Cambridge University Press.
- McCarthy, M., & Carter, R., (1997). *Language as discourse. Perspectives for language teaching*. London, England: Longman.
- McDowell, C. (1995). Assessing the language proficiency of overseas-qualified teachers: The English Language Assessment (ELSA). In G. Brindley (Ed.), *Language assessment in action* (pp. 11-29). Sydney, Australia: National Centre for English Language Teaching and Research, Macquarie University.
- McNamara, T. (1996). *Measuring second language performance*. London, England: Longman.
- McNamara, T. (1997). Interaction in second language performance assessment: Whose performance? *Applied Linguistics*, 8, 446-466.
- Nickerson, C. (2010). English as a key resource to business and development. In G. Forey & J. Lockwood (Eds.), *Globalization, communication and the workplace* (pp. 25-41). New York, NY: Continuum International.
- O'Neil, D. A. (2003). Offshore outsourcing takes hold. *Mortgage Banking*, 64(3), 38-43.

- Pal, M., & Buzzanell, P. (2008). The Indian call center experience: A case study of changing discourses of identity, identification, and career in a global context. *Journal of Business Communication*, 45, 31-60.
- Pienemann, M., & Johnston, M. (1987). Factors affecting the development of language proficiency. In D. Nunan (Ed.), *Applying second language acquisition research* (pp. 45-141). Adelaide, Australia: National Curriculum Resource Centre.
- Reeves, N., & Wright, C. (1996). *Linguistic auditing: A guide to identifying foreign language communication needs in corporations*. Bristol, England: Multilingual Matters.
- Revell, R. (1994.). From control to management: Practices and trends in English for occupational purposes. In J. Boyle & P. Falvey (Eds.), *English language testing in Hong Kong* (pp. 300-313). Hong Kong: Chinese University Press.
- Schegloff, E. A. (1982). Discourse as an international achievement: Some use of "uh huh" and other things that come between sentences. In D. Tannen (Ed.), *Analysing discourse: Text and talk* (pp. 71-93). Washington, DC: Georgetown University Press.
- Weir, C. (1998). *Understanding and developing language tests*. New York, NY: Prentice Hall.

Bio

Jane Lockwood is currently Head of the English Language Centre and concurrently Associate Professor in the Department of English at City University of Hong Kong. She has lived and worked in Asia for the last 15 years and carries out regular research in the business processing outsourcing (BPO) industry in the Philippines and India.