# Revisiting Communicative Competence in the Teaching and Assessment of Language for Specific Purposes

SHONA WHYTE [a]

*[a] Université Côte d'Azur, CNRS, BCL, FRANCE*
Email: whyte@unice.fr

## Abstract

The term *communicative competence* captures the notion that the ability to use language in interaction requires not just control of linguistic form but also awareness of rules of use in different contexts (Hymes, 1972). Communicative competence is a slippery term: different actors in second language (L2) research, education, and assessment interpret the term in a variety of ways and use it for a range of purposes, perhaps particularly in the field of languages for specific purposes (LSP). This is unfortunate because it is a key concept in LSP, as in applied linguistics more generally. Communicative competence can be considered to be the target of second language acquisition, a main goal of second or foreign language teaching and learning, or the object language testers seek to measure via performance tests. In addition, current interpretations of communicative competence may be somewhat questionable adaptations of Hymes' concept, modified and often simplified to reflect current approaches in both formal and functional linguistics, and to respond to practical concerns in language teaching and testing. This paper seeks to re-examine communicative competence from three perspectives - L2 research, teaching, and testing - highlighting problems in terms of theory and practice with respect to LSP. Drawing on recent research on indigenous assessment criteria, the paper concludes with a revised model of communicative competence for LSP, offering a richer interpretation closer to the original concept and to current concerns in the field.

**Keywords:** Language for specific purposes (LSP), communicative competence, language education, second language (L2) research, second and foreign language teaching

## Introduction

Communicative competence (Hymes, 1972) is one of the most important notions in languages for specific purposes (LSP) teaching and learning. Many key texts in LSP focus on language users'

abilities to communicate effectively, or simply 'get things done' in particular contexts of communication. LSP is related to "the communicative needs of speakers of a second language in facing a particular workplace, academic, or professional context" (Basturkmen & Elder, 2004, p. 672), and these needs include "not only linguistic knowledge but also background knowledge relevant to the communicative context in which learners need to operate" (Douglas, 2013, p. 371). English for Specific Purposes (ESP) focuses on "the demands placed by academic or workplace contexts on communicative behaviours" (Hyland, 2002, p. 386) and "the language, skills, and genres appropriate to the specific activities the learners need to carry out in English" (Johns, 2013, p. 2). The importance of contexts and goals for communication is clear, creating a natural connection between LSP and the notion of communicative competence.

Yet as the notion has evolved over time, different subfields of applied linguistics like second language (L2) research, teaching, and testing have pursued divergent interpretations, creating contradictions for LSP, which has traditionally drawn on these subfields. This paper re-examines communicative competence from these three perspectives to highlight tensions between theory and practice in LSP and propose a revised model which constitutes a more faithful representation of Hymes' original notion and is also closer to current concerns in LSP assessment.

## The origins of the term communicative competence

This concept was first proposed by Hymes (1972) in an essay where the sociolinguist argued for a linguistic theory which could focus on "the capacities of persons, the organisation of verbal means for socially defined purposes, and the sensitivity of rules to situations" (p. 292). Hymes was reacting to Chomsky's famous distinction between the **competence** of "an ideal speaker-listener, in a completely homogeneous speech community, who knows its language perfectly," on one hand, and "errors (random or characteristic) in applying his knowledge of the language in actual **performance,**" on the other (Chomsky, 1965, p. 3). Hymes (1972) recognised this distinction as a contemporary interpretation of a tradition leading back to Saussure and even Humboldt, and questioned the prioritisation of linguistic competence, that is, "tacit knowledge of language structure" (p. 271) over performance, or "imperfect manifestation of underlying system" (p. 272).

Hymes saw the restrictive view taken by Chomskyan linguistic theory as "almost a declaration of irrelevance" of sociolinguistics (p. 270), and one which "omits almost everything of sociocultural significance" (p. 280). Hymes sought to rehabilitate a sociolinguistic interest in rules of use, since these, he argued, "are not a late grafting" in child language acquisition processes (p. 279), but are instead acquired at the same time as structural knowledge. He pointed out that even Chomsky admitted "the possibility of stylistic 'rules of performance'" (p. 280): since rules imply competence and thus contradict the competence/performance dichotomy, Hymes took up the challenge of modelling what he termed communicative competence.

This richer conception of competence includes four types of knowledge together with an "ability for use" which is related to each of the four dimensions. These are shown in Table 1, whose wording derives from Hymes' (1972) text.

For Hymes, communicative competence thus includes speakers' knowledge of linguistic and sociolinguistic rules as well as their ability to use this knowledge in interaction. It is distinct from actual language use in interaction, which depends not only on speakers but also their interlocutors and unfolding events, and comes under the heading of performance. This view thus calls into question Chomsky's competence/performance distinction between linguistic knowledge and language use. Where Chomsky set up a binary opposition, Hymes proposed three categories covering speakers'

knowledge of language rules, their ability to use rules to interact, and actual language use during events involving others. While the notion has been applied to both written and spoken language, it is clear from Table 1 that Hymes was thinking primarily of oral interaction, and this is also the focus of the work reviewed in this paper.

**Table 1** *Dimensions of Communicative Competence* (Hymes, 1972)

| COMMUNICATIVE COMPETENCE | | PERFORMANCE |
|---|---|---|
| Knowledge | Ability for use | Actual use & events |
| 1  *What is possible*<br>Systemic possibility<br>"Grammaticality" (in terms of syntax but also culture, communication)<br>2  *What is feasible*<br>Psycholinguistic reality<br>Constraints on memory, perception<br>3  *What is appropriate*<br>Situational judgement<br>Acceptability in context<br>4  *What is performed*<br>Actual occurrences | • Motivation<br>• Affective & volitive factors<br>• Capacities in interaction (e.g., composure, presence of mind, stage confidence) | • Behavioural record<br>• Imperfect or partial realization of individual competence<br>• Interaction between individual competence, competence of others, and properties of events |

In the five decades since its initial formulation, the concept of communicative competence has evolved in different directions in different areas of applied linguistics. These include work on genre theory and academic literacy[1], for example, which focus on written language and are no doubt less central to our concerns with communicative competence in L2 education and assessment. The field of LSP has historical ties with three areas of applied linguistics which are arguably of most relevance here. Second language (L2) research, a relatively young discipline usually dated to Corder (1967) and Selinker (1972), has traditionally often adopted an LSP perspective (e.g., Selinker & Douglas, 1985). LSP is also demonstrably a practitioner-led field with particular interest in addressing issues of teaching and learning (Johns, 2013). The field of language testing, too, is commonly involved in LSP due to the importance of language tests in many forms of institutional gatekeeping (Fulcher, 2013). How has the term communicative competence influenced these different disciplines?

This paper revisits the notion of communicative competence in each of these fields in turn, highlighting inconsistencies, shortcomings, and conflicts in ways that are hopefully helpful to the field. It concludes with a revised model of communicative competence in LSP, drawing on recent research on indigenous assessment criteria.

## Second language research perspectives on communicative competence

Like general linguistics, research in second language acquisition (SLA) has long maintained an often useful distinction between an idealised, abstract, subconscious knowledge of a language, on one hand, and the messy, error-ridden, or elliptical language use or behaviour, on the other. Formal linguistics has traditionally focused on competence, viewed as an abstract system of syntactic rules underpinning universal grammar, for instance. Functional approaches, on the other hand, may pay more attention to performance as a source of rules of language use. L2 research initially took a formalist route.

---

[1] Readers interested in this research are referred to Bhatia (2014), Dressen-Hamouda (2012), Hyland (2006), and Paltridge (2012) for useful overviews of LSP dimensions of genre theory, literacy, or ethnographic approaches including language socialisation and indexicality.

**Early interlanguage research**

The notion of competence in L2 research comes from understanding that learner language is more than the sum of its parts, not learned by piecing together words according to rules, but rather consisting in a subconscious, abstract system which informs real-time language processing. It is different from performance, which contains mistakes due to processing constraints such as memory. Corder (1967) was the first to note the systematic nature of L2 learner errors, and to view this as "evidence that the learner uses a definite system of learning at every point in his development" (Corder, 1967). Corder's L2 "system, or 'built-in' syllabus" led Selinker (1972) to coin the term **interlanguage** as "a separate linguistic system" resulting from "the learner's attempted production of a target language norm." L2 researchers took up the challenge of describing this system, in terms of divergence from native-speaker norms, and with respect to development in linguistic accuracy, complexity and fluency over time (Skehan, 2009). Originally under the banner of Chomskyan generative SLA, this cognitivist approach to interlanguage research has been perhaps most forcefully defended by Kevin Gregg with emphasis on "three key words: *explanation,* not description or prediction; *acquisition,* not use; *competence,* not behaviour" (Gregg, 1990, p. 365). Gregg rejects variationist approaches to L2 research (Tarone, 1983), sociocultural theory (Lantolf, 2000), and complex dynamic systems theory (Larsen-Freeman & Cameron, 2008) on the same basis, reiterating an exclusive focus on the "linguistic competence(s) of an individual – the standard view in theoretical linguistics" (Gregg, 2010, p. 552-3) and dismissing "what everyone likes to call 'communicative competence'" out of hand (Gregg, 1990, p. 365).

Other cognitive linguists disagree: for Lakoff and Johnson (1999), there is no Chomskyan person, for whom language is pure syntax, pure form insulated from and independent of all meaning, context, perception, emotion, memory, attention, action, and the dynamic nature of communication.

Views which take more account of communicative competence in L2 development, of particular importance in LSP, are considered in the next section.

**Theory underpinning communicative language teaching**

As noted in the introduction, sociolinguists and applied linguists have historically taken an wider view of the domain of language study. Hymes famously claimed that "there are rules of use without which the rules of grammar would be useless," though his own work focused on sociolinguistic aspects of first language use. An early interpretation for L2 research of these wider dimensions of communicative competence was proposed by Canale and Swain (1980). Concerning the principles of communicative language teaching (CLT), these authors recommend giving priority to opportunities for "meaningful communicative interaction" in order to provide learners with "the information, practice and much of the experience needed to meet their communicative needs in the second language" (Canale & Swain, 1980, p. 27-8). Their model is shown in Table 2, which uses the 1980 text with slight adaptation for clarity of exposition.

As noted earlier, Hymes revised the Chomskyan conception of competence by a) expanding the notion of linguistic competence to include knowledge of sociolinguistic appropriateness and b) adding an ability for language use corresponding to Hymes' capacity to interact  (Table 1). Table 2 shows that Canale and Swain's model also includes these two types of knowledge – grammatical and sociolinguistic -  but it assigns sociolinguistic knowledge to the category of "use." Their model also adapts part of Hymes' "ability for use" to apply specifically to L2 users: they propose a third dimension of communicative competence consisting in compensatory strategies, which are used when communication breaks down, notably because competence related to knowledge and/or use is lacking. These strategies may be non-verbal, or may involve paraphrasing or other avoidance moves (and some are common to the performance failures of L1 users).

**Table 2** *Canale & Swain's Model (1980)*

| COMMUNICATIVE COMPETENCE | | |
|---|---|---|
| **Compensatory strategies** | **Knowledge**<br>*Grammatical competence* | **Use**<br>*Sociolinguistic competence*<br>• sociocultural rules<br>• rules of discourse |
| | Knowledge of<br>• lexical items<br>• rules of<br>    ○ morphology<br>    ○ syntax<br>    ○ sentence-grammar<br>    ○ semantics<br>    ○ phonology | *Sociocultural rules* of use will specify the ways in which utterances are produced and understood appropriately with respect to components of communicative events<br><br>*Rules of discourse*.<br>The focus is the combination of utterances and communicative functions (not grammatical well-formedness nor sociocultural appropriateness) |
| **Strategic competence**<br>• verbal and non-verbal communication strategies<br>• for communication breakdowns (due to performance variables, insufficient competence) | Strategies that relate primarily to *grammatical competence* (e.g., how to paraphrase grammatical forms that one has not mastered or cannot recall momentarily) | Strategies that relate more to *sociolinguistic competence* (e.g., various role-playing strategies, how to address strangers when unsure of their social status) |

This model gives grammatical knowledge greater prominence than Hymes did, since here knowledge is restricted to the main categories of linguistic analysis (phonology, syntax, semantics), and seen as separate from sociolinguistic awareness. Critics like Widdowson (2017) have contested the priority accorded this kind of knowledge of the linguistic code in instrumentalisations of communicative competence in CLT. In real-world contexts, he argues:

> appropriateness is determined by variable contextual factors, and so communicative function is not inscribed in particular encoded forms. To suppose otherwise is to confuse the semantics of the language code with the pragmatics of its use, and so to misrepresent the very nature of communication.

This objection is at least partially answered in Canale and Swain's model, which in addition to knowledge and use, provides for "compensatory strategies." Strategic competence helps speakers to circumvent problems caused by insufficient mastery of grammar. The authors' formulation implies the existence of a final state where such strategies are no longer needed, and the authors suggest

> knowledge of how to use such strategies may be particularly helpful at the beginning stages of language learning, and it is to be expected that the need for certain strategies may change as a function of age and second language proficiency.
> (Canale & Swain, 1980, p. 31)

The possibility that some kinds of strategic ability may remain important for all L2 users is left open, and since this point is particularly important to LSP, we will return to it in discussion of native speaker (NS) norms.

**Instructed second language acquisition**

Since much L2 research is conducted in classroom contexts, and pedagogical implications are frequently drawn from its findings, the subfield of instructed second language acquisition (ISLA) is receiving growing attention (Long, 2017). Long defines ISLA as language learning *"when the learning processes are influenced, or at least intended to be influenced, by teachers, classmates, or pedagogic materials"* (p. 8). The objective is to understand how "different kinds of intervention" exert "objectively measurable effects on interlanguage development" and on "learners' ability to perform real-world tasks" (p. 9). The performance of tasks is given new theoretical importance: "whereas interlanguage development has traditionally been viewed as leading to improved communicative abilities," the reverse is now thought to be true: "it is improvement in the ability to perform progressively more complex communicative tasks that drives language learning" (p. 9).

The field draws on L2 theory as well as empirical findings to motivate research and the main goal is to identify causal relationships between language teaching and learning in order to improve L2 learning or teaching. Long supports a cognitive-interactionist theory of SLA which informs communicative instructional programmes "such as immersion, TBLT, and CLIL" and, as noted, views communicative language use as an important trigger for L2 development.  In such programmes, teaching and learning focus on "the non-linguistic syllabus, with the L2 in theory learned incidentally through being used communicatively as the medium, not the object, of instruction." Long (2017) argues against "explicit instruction and a focus on language as object" since this kind of teaching disrupts the learning of "crucial non-linguistic syllabus content," and suggests that *"focus on form*, with its temporary brief switches to intentional language learning during otherwise communicative lessons, is a major improvement in this regard." (p. 25) Long goes on to argue that SLA researchers agree on the central position of "incidental and implicit L2 learning in adults" but that this is "still a minority position in the world of language teaching" (p. 23). His approach also faces criticism from L2 researchers outside instructed contexts.

**Sociocultural and intercultural approaches**

As suggested earlier, Widdowson is not the only critic of SLA models of interlanguage development and definitions of communicative competence. In a seminal paper marking what has been called a 'sociocultural turn' in L2 research, Firth and Wagner (1997) also contest what they see as a consistently reductive view of L2 communicative success. These authors reject SLA's emphasis on "the foreign learner's linguistic deficiencies and communicative problems" (p. 288), and resist the suggestion that the L2 user is "in a phase of transition" (p. 292), to be viewed as a "deficient communicator struggling to overcome an underdeveloped L2 competence, striving to reach the 'target' competence of an idealised NS" (p. 295-6).

Firth and Wagner (1997) thus express concerns similar to Widdowson's (2017) with respect to a preoccupation with language accuracy in L2 research and default reference to NS norms in studies of interlanguage development. Outside the language classroom, it is of course obvious that much effective communication is successfully undertaken by interlocutors who are neither native speakers (L1 users) nor language learners, but rather L2 users whose interlanguage is stable yet not nativelike. Once again, this argument is particularly relevant to LSP and we return to this population in the third section of the paper, after examining the classroom context, this time from the perspective of language teachers.

## Second language teaching: Common European Reference Framework (CEFR)

The field of second and foreign language education both predates L2 research (McLelland & Smith, 2014; Smith & McLelland, 2018) and takes a wider perspective (Kramsch, 2000; Spada, 2013). It has been dominated in recent decades by communicative approaches, often justified with reference to SLA theory. In a paper which documents inappropriate applications of SLA findings to classroom practice, Spada (2015) acknowledges the influence of Hymes while also denouncing the excesses of strong CLT in effecting the "pendulum swing that took place in L2 teaching in the late 1970s and early 1980s"(p. 4). Also drawing heavily on the concept of communicative competence, a key development in language education in the past twenty years involves competence-based frameworks for evaluation. I focus on the extremely influential Common European Reference Framework for Languages (CEFR, Council of Europe, 2001), which Widdowson (2017) sees as "the functional equivalent to the formalist concept of interlanguage."

### Notions and communicative functions: the CEFR

A cross-linguistic competency framework based on fine-grained 'can do' statements, the CEFR shares with interlanguage research a concern to identify stages of approximation of native-speaker competence. Since its inception, much work has been devoted to refining the six level descriptors of the CEFR and its success has been ascribed to
its combination of what is familiar (the traditional distinction between 'beginner', 'intermediate', and 'advanced' levels) and what is new (an elaborate system of descriptors giving communicative content to the levels beginner/basic, intermediate/independent, and advanced/proficient). (Hulstijn, 2007, p. 663)

The CEFR grew out of notional-functional approaches dedicated to communicative goals, based on the conviction that "what people want to do through language is more important than mastery of language as an unapplied system" (Wilkins, 1973, p. 136-7). Linguistic knowledge is not "an end in itself," rather, the goal of CLT is "the ability to use language, to do with language the kinds of things ones needs or wants to do with it" (Van Ek & Alexander, 1975, p. 19). Wilkins provided early groundwork on what he termed a "situational syllabus" to rationalise language teaching by addressing "first what is grammatically necessary" and "secondly what constitutes a speaker's communicative competence" (Wilkins, 1973, p. 143). His paper lists notional categories (time, quantity, space, matter, case, and deixis) and categories of communicative function (modality, moral discipline, suasion, argument, rational enquiry, and personal/emotional/interpersonal emotions). This work informed the first formulation of the CEFR, which is explicitly predicated on action-based CLT and thus important for LSP teaching and learning.

### Empirical and theoretical justifications of the CEFR

Although a supporter of the humanist ideals behind its conception, Hulstijn (2007; 2014) points out that the CEFR is built on rather "shaky ground" both in empirical and theoretical terms. He notes that "its empirical base consists of judgments of language teachers and other experts with respect to the scaling of descriptors" and that "the CEFR scales lack empirical support of what L2 specific knowledge and skill is minimally required for performance considered adequate in terms of communicative functioning" (Hulstijn, 2007, p. 665-6). He calls for "empirical support based on performance data of L2 learners" (Hulstijn, 2014, p.16) but also queries the theoretical underpinning of the framework: "we do need to know first what language proficiency means in the case of NSs before we can consider the case of NNSs" (Hulstijn, 2007, p. 664).

**CEFR as standard for L2 proficiency testing**

Considering only the empirical side of these criticisms, recent work in relation to the CEFR has tackled questions of the reliability and validity of CEFR-based tests, and proficiency correlates of CEFR levels for specific languages. Deygers, Van Gorp, and Demeester (2018) compared the interpretation and operationalisation of CEFR descriptors in two tests of oral proficiency in Dutch which are used to certify B2 proficiency for international student admission to university in Flanders (Belgium). The two tests, the ITNA (created by a Flemish consortium of university language centres) and the STRT (an international test developed by the Dutch Language Union), employ similar speaking tasks: a 25-minute oral interview with a trained examiner, involving a presentation based on graphs or tables and an argumentation task. Candidates are scored on five linguistic criteria (vocabulary, grammar, coherence, pronunciation, and fluency) by two raters using A2, B1, B2, and C1 band descriptors. The authors note differences in rating conditions: the ITNA tests are scored immediately after the test, while STRT are recorded. Rater profiles also differ across the two tests: ITNA examiners are generally experienced Dutch L2 teachers who train and test several times a year, while STRT raters tend to be younger students of linguistics or communication who have initial training including a single trial session.

The researchers compared the scores of 82 students on both tests (taken one week apart, STRT first) and found significantly lower scores on the ITNA. Examination of the criteria used in each test found departures from the CEFR wording, and thus limited overlap between descriptors in the two tests. The remainder of the study involved statistical tests of components of oral scores obtained by students on each exam. Detailed comparison of the five scoring criteria which were common to the two tests revealed that

> there is a consistent significant difference between the probability of attaining a score of at least B2 on the ITNA or one of the STRT tasks (p < .05). This indicates that the B2 threshold is interpreted or operationalized differently on the STRT and on the ITNA test. (Deygers *et al.*, 2018, p. 9)

The authors conclude that "this study has yielded no data to indicate that corresponding CEFR-based criteria used to measure the same candidates in near-identical tasks can be considered equivalent" (Deygers *et al.*, 2018, p. 12) and that therefore "the CEFR may be a useful inspiration for test developers to reflect on language proficiency levels, but it is not a standard that can simply be applied to reach equivalent scores" (Deygers *et al.,* 2018, p. 13).

In the limitations section of the paper, the authors cite the "real-life" context of data collection as a drawback, but do not consider a number of other potential confounding variables. The first is a rater effect – the possibility that the more experienced ITNA examiners grading in situ were more strict than their younger STRT counterparts working from recordings. The authors also note differences in the populations taking the two exams: two thirds of ITNA candidates were potential students, compared to just over half of STRT candidates, more of whom were professionals (who are often held to a lower standard). The third concern is for an order effect – candidates who enrolled for ITNA were offered STRT as an "extra opportunity" for practice which they may have approached in a different manner from their preferred choice of exam.[2] Analysis of actual data from exam interviews would be useful to pursue this speculation. Nevertheless, the finding that a given L2 user was significantly more likely to reach the cut-off score of B2 on one test compared with another poses serious problems for the CEFR

---

[2] The researchers did not control for order of test administration because ITNA scores are communicated on the day of testing and it was anticipated that successful candidates would not agree to take the second test (STRT). ITNA candidates were therefore approached and offered an STRT first, since these scores are not published immediately.

as a testing standard.

**Criterial features in the English Profile Project**

Another attempt to shore up the empirical foundations of the CEFR is the Language Profile Project, which aims to produce "reference level descriptors" specific to each national language (Council of Europe, 2005). The English Language Profile (EPP) involves the analysis of learner corpora compiled from Cambridge test data in order to identify "criterial features" characteristic of learner performance in L2 English at different CEFR levels. In this project, researchers assume that "in addition to whether a learner fulfils the communicative functions required by the task," it is possible to identify "certain linguistic properties that are characteristic and indicative of L2 proficiency at each level, on the basis of which examiners make their practical assessments" (Hawkins & Filopovic, 2012). One outcome of the EPP is the English Vocabulary Profile and another the English Grammar Profile (EGP). Both aim to relate the general CEFR level descriptors to specific features of competence in English L2. The latter is described in some detail by O'Keeffe & Mark (2017).

The authors worked with a CEFR-calibrated learner corpus culled from Cambridge written exams taken over 13 years (1999-2012) by learners with 143 first languages. Some 55 million words (64 million tokens from 267 000 passing scripts) were annotated using the corpus tool Sketch Engine; the British National Corpus (written) was used for comparison where necessary. In keeping with the overall project goal of developing practical tools for teachers and learners, the EGP search inventory was based on what the authors refer to as the "ELT canon" or established approach to English language teaching apparent in textbooks and discussion with teachers. O'Keeffe and Mark accordingly searched the corpus for examples of language use in 19 superordinate grammatical categories (e.g., adjectives, negation, present time) using a criteria-based approach. To be considered characteristic of a certain CEFR level, a form must meet frequency, accuracy, and dispersion criteria. Table 3 shows how the authors operationalised these criteria:

**Table 3** *English Grammar Profile Criteria (O'Keeffe & Mark 2017)*

| Criterion | Description | Benchmark |
|---|---|---|
| Frequency | Is there sufficient evidence of a structure at this level? | Frequencies lower than BNC written frequency were not considered |
| Accuracy | Is there an adequate rate of correct uses? | 60% correctness in both linguistic and pragmatic terms |
| Dispersion | Is usage distributed across a range of | |
| | Users | Standard deviation to measure dispersion across users |
| | L1 families | No more than 30% of use from one language family |
| | Contexts of use | Range of styles, formats, and registers |
| | Tasks | Cross-check against task instructions and constraints |

O'Keefe and Mark (2017) applied these criteria iteratively for each grammatical form at each CEFR level (using pass scripts to ensure correlation with examiners' judgements), writing "can-do" statements "to represent the use of a grammatical item with a particular form and/or use, at a given level" and checking for other uses of the same form (p. 471). In keeping with their focus on "the development of grammar competence" (p. 476) as opposed to error analysis or fossilisation, their work highlights the increasing lexicogrammatical complexity and pragmatic subtlety of learner production at higher levels. They show, for example, that a pattern involving a pronoun followed by a linking verb, optional adverb, adjective and *that* clause is instantiated in increasingly sophisticated ways from lower

to higher CEFR levels (p. 478): *I am sure* (A2 Norwegian), *it seems obvious that* (B2 French), *it is highly unlikely that* (C1 Russian). In contrast to claims about a ceiling effect where learners no longer progress (Thewissen, 2013), these authors discern "greater complexity of meaning" and "greater dexterity of use" with advancing levels (p. 478), including pragmatic development. The study is an example of a corpus linguistics approach to learner data which is perhaps more compatible with traditional L2 research than other CEFR-related developments (e.g., Bardovi-Harlig, 2000).

This section has shown that the interpretation and operationalisation of communicative competence in the dominant language teaching framework in Europe today is somewhat lacking in theoretical underpinning, difficult to operationalise, and strongly oriented toward linguistic competence, raising questions for its application to LSP contexts. What about language for specific purposes testing?

## Language for specific purposes testing

Much has been written about the challenges of designing communicatively appropriate and effective tests of languages in specific purposes contexts (Douglas, 2001; Harding, 2014; Elder, McNamara, Kim, Pill, & Sato, 2017). LSP testing is justified by "the need to describe minimum levels of performance for work in high-stakes areas, such as speaking in air traffic control;" L2 users must therefore be able to "communicate efficiently in contexts where a failure to do so would put others in danger" (Fulcher, 2013, p. 14). Many would agree with Lockwood (2012) that professional communication is often "still very underresearched" (p. 23). In his book on language testing, Fulcher (2013) also addresses the use of tests for perhaps more controversial gatekeeping functions, stressing the burden on language testers to consider the validity of "all possible uses" of their tests (p. 20). With this in mind, a number of recent studies of LSP testing have questioned the validity of performance tests used in a range of professional contexts. A recent paper by Elder and her colleagues (2017) considers LSP testing research in medicine, veterinary science, air traffic control, as well as academic settings such as scientific research presentations and university entrance tests. Douglas (2001) has argued that while the language content and method of LSP tests (i.e., test tasks) are "fairly well understood" (p. 173), the same cannot be said of assessment criteria. The following studies of LSP tests focus on what has been termed **indigenous** criteria, that is, the views of occupational experts, non-language specialists, or linguistic laypersons. LSP studies of such assessment criteria are reviewed in the following subsections.

### Naturally occurring scientific communication and professional language tests

One of the first studies to take an indigenous perspective on LSP use (Jacoby, 1999; Jacoby & McNamara, 1999) involved the observation of research physicists at an American university during regular lab meetings. With the goal of characterising academic talk involving both L1 and L2 speakers, Jacoby used conversation analysis and grounded theory to analyse these speech events, which typically involved a conference paper rehearsal and ensuing feedback. She found that the focus of the group's attention was exclusively on content, that is, the effective presentation of scientific material. The criteria used by these researchers to evaluate presentations were timing, newsworthiness, visual coherence,[3] clarity, economy of expression, argumentation, content accuracy, technical delivery, and overall quality. Only "a tiny subset of comments" concerned L2 users' errors (spelling, prepositions, irregular past), and then only when these occurred in written presentation material (Jacoby & McNamara, 2013, p. 232). The authors argue:

> It may be that certain problems of clarity, delivery, and economy, etc. addressed in feedback to non-native-speaker presenters could be traced to the fact that English is not their first language,

---

[3] Data were collected before the generalisation of Powerpoint, and the scientists used overhead transparencies.

> but what is significant is that the comment givers do not treat these problems as such during indigenous assessment interaction. (p. 233)

They contrast these indigenous criteria with assessment criteria typical of LSP tests, taking the example of an Occupational English Test (OET) which is used to screen international health professionals for entry to Australia. The authors highlight serious problems with this medical English test. By focusing on native speaker norms, the OET gives undue emphasis to linguistic competence and so "foregrounds what for the physicists is a marginal issue in communicative success" (p. 234). Although the test purports to assess specific purpose communicative competence, it uses standard language assessment criteria based on the four skills, without reference to the content-related, interactional dimensions which were shown to be so important in the physicists' assessments of effective communication.

One might object that the English requirements for research physicists and medical practitioners are likely to be somewhat different, yet closer analysis reveals a number of communicative functions which are common to the two contexts. Further research on the OET was conducted by Pill (2013), also reported in Elder *et al.* (2017). The impetus for Pill's (2013) study came from concerns among medical professionals about the validity of the OET, specifically that it was failing to identify those international medical graduates (IMG) who were best able to benefit from preparation courses for professional certification to practice medicine in Australia. Like Jacoby's study, this research involved the gathering of opinions on specialised communicative competence from domain experts as opposed to language professionals. This researcher's data included medical educators' commentary on video recordings of consultation scenarios with simulated patients, as well as their actual written feedback to trainees in real medical consultations. Based on the comments of these experts, Pill concluded that the four existing OET criteria – intelligibility, fluency, appropriateness of use, and resources of grammar and expression – while relevant, nevertheless represented a "somewhat restrictive view of language as a decontextualized set of elements" (Elder *et al.*, 2017, p. 18). The medical experts noted that IMGs lacked pragmatic awareness, for example, asking directly "Do you want to harm yourself?" instead of the more circumspect "Sometimes when people feel down, they feel like escaping/hurting themselves. Do you ever feel like that?"(p. 18). They also identified failings in strategic competence: IMGs were found to be "scared of open questions because they think they'll lose time […] it always works the other way round"(p. 18). The inclusion of two new criteria in the OET to reflect these points - *clinician engagement* and *management of interaction* – now allows the test to cover a wider interpretation of interactional competence thus increasing validity. Elder and her colleagues interpret this study as an example of real-world consequences of LSP testing: where the views of applied linguists alone produced an original OET based on narrow linguistic criteria which resulted in poor admission decisions, a revised test including the indigenous criteria arising from Pill's (2013) research seems likely to prove a more valid indicator of LSP communicative competence.

**Professional views of assessment criteria in other contexts**

A similar comparison of the views of domain specialists compared with applied linguists was conducted in a veterinary medicine programme at Iowa State University (Douglas & Myers, 2000), this time focusing on apprentices communicating in their L1. Again drawing on a real-world problem of assessment of communicative competence, this research focused on the evaluation of a special Rural and General Practice option in which students were assessed via an Individual Process Assessment (IPA), where they role-played interactions with clients (animal owners). Douglas and Myers (2000) examined evaluations of video recordings of student performances in this context by veterinary professionals, veterinary students, and applied linguists, and compared the categories used in commentary by each of three groups. While the researchers found extensive overlap across groups, and with the 16 official IPA criteria, they also identified differences in emphasis, and a number of gaps. The most explicit comments were furnished by the veterinary experts, and only this group used all 16

categories in the IPA. The study also identified six additional categories absent from the IPA criteria but mentioned by all three groups: *demeanor, knowledge base, timing, coverage, phraseology*, and *appearance* (Douglas & Myers, 2000, p.74-5). The researchers suggest that some, but not all of the new criteria might be included in a new IPA (e.g., *knowledge base* but not *appearance*); all these criteria are included in Table 4 for comparison with the other research reviewed here.

A further example of research into indigenous assessment criteria comes from the field of civil aviation in Korea (Kim, 2012; Kim & Billington, 2016). As with the OET described above, the Korean English Proficiency Test for Aviation (EPTA) was perceived as a poor indicator of the professional competence of pilots and air traffic controllers, and as in the previous studies, the researchers investigated professional views of actual test data, here authentic audio recordings of pilot-controller communication. Kim (2012) elicited the views of 3 experienced pilots and 5 controllers of 6 episodes recorded during non-routine, abnormal, emergency, and distress situations with the goals of a) identifying specialist language and important communicative practices, and b) relating these to official EPTA assessment criteria: *comprehension*, *fluency*, *interactions*, *pronunciation*, *structure*, and *vocabulary*. The researchers found discrepancies between professional and language test criteria. Their professional informants argued that since this type of communication is highly codified, performed by multitasking individuals in high-pressure conditions, assessment criteria based on fluency and (grammatical) structure were irrelevant if not counterproductive.

Another part of the study analysed transcripts of the actual pilot-controller communication in these unusual incidents and found that communication errors were not exclusively due to L1-influenced pronunciation or other errors by L2 speakers. Instead, responsibility for communication problems was shared by both L1 and L2 speakers; indeed Kim and Billington (2016) suggest that native speakers "are just as likely to encounter difficulty with unfamiliar accents, and should be held to the same standards for both Pronunciation and Comprehension" (p. 20). In this instance there have been no moves to bring the official language test closer to indigenous criteria, and Korean authorities are currently circumventing negative effects on experienced staff by publishing answers in advance. Elder and her colleagues (2017: p. 19) claim that by insisting on a narrow definition of communicative competence, which is restricted essentially to linguistic competence, such exams will continue to exclude individuals from participation in professional settings who may in fact be competent to practise, and to allow others access to professional practice
whose actual (in)competence may cause problems of communication.

**Academic English and lay assessment of language proficiency**

The final study in this overview of attempts to determine indigenous assessment criteria concerns this time not professional practice, but university entrance, bringing us to the domain of English for Academic Purposes (EAP). Sato (2014) researched the assessment of oral competence as measured in the Chinese College English Test-Spoken English Test (CET-SET), a monologue, and in the oral component of the Cambridge English exam, a paired interaction. Investigating a hypothesised mismatch between linguist and lay interpretations of oral competence, Sato (2014) collected the evaluations of 23 non-linguist L1 and L2 English speakers of 13 test recordings (from 7 CET-SET and 3 Cambridge English tests) and compared them with official ratings by trained evaluators. These lay judges were asked to assign a holistic score from 1 to 7, then justify their grades in stimulated recall and follow-up interview.

The study found similarities and differences between lay and professional judgements with respect to monologue and paired tasks, and also highlighted discrepancies between the criteria used by lay judges and language testers. The lay judges mentioned *demeanour*, *non-verbal behaviour*, *pronunciation*,

**Table 4** *Indigenous and Linguistic Criteria in LSP*

| Criteria | Everyday communication<br>Sato 2013 | Veterinary science<br>Douglas & Myers 2000 | Research physics<br>Jacoby 1999, Jacoby & McNamara 1999 | Clinical medicine<br>Pill 2013 | Civil aviation<br>Kim 2012, Kim & Billington 2016 |
|---|---|---|---|---|---|
| INDIGENOUS | | | | | |
| general impression | **overall impression** (ability, performance, message, proficiency) | | overall quality | | |
| | demeanor, non-verbal behaviour | **demeanor, appearance** | technical delivery, visual coherence | | |
| content | **content** | **knowledge base, coverage**, admitting lack of knowledge | newsworthiness, content accuracy | | (lack of) **professional knowledge** |
| | | questioning skills (type, duplication, summarizing, jargon, verification) | | | |
| discourse strategies | | organisation, introduction, closing the interview, transitional statements | clarity, economy of expression, argumentation | **management of interaction** | |
| interactional awareness | | client understanding, establishing rapport, empathetic statement, tone of voice, challenging the client | | **clinician engagement** | **accommodation strategies** |
| | | **timing** | timing | | |
| LINGUISTIC | | | | | |
| interaction | **interaction** | | | appropriateness of use | interactions |
| fluency | **fluency** | | | fluency | *fluency* |
| intelligibility | pronunciation | | | intelligibility | pronunciation |
| receptive skills | | | | | comprehension |
| grammar & vocabulary | linguistic resources (vocabulary, grammar) | **phraseology** | grammar (spelling, prepositions, irregular past) | resources of grammar and expression | *structure* |
| | | | | | vocabulary |

*linguistic resources (vocabulary, grammar)*, *fluency*, *content*, *interaction*, and *overall impression (ability*, *performance*, *message*, *proficiency*) in justifying their assessments. The most frequently mentioned category was *overall impression*, particularly *proficiency* (20% of comments across both

test types), followed by *content* (15%). For the paired speaking test, *interaction* came next (12%), while *fluency* ranked higher in the monologue task (13%). Both were more frequently mentioned than *linguistic resources* (11%; Sato, 2014, p. 137). These results suggest that to gauge communicative effectiveness, ordinary people believe they rely more on global perceptions of ability and on the content of a speaker's contribution than on linguistic competence.

Table 4 compares the indigenous variables and exam criteria used in the different content domains discussed in this section. The indigenous criteria generally relate to a broad conception of communicative competence including extralinguistic variables such as *appearance* and *demeanor*, and *content knowledge*, while the linguistic criteria refer to a narrow range of language skills (comprehension) and subskills (pronunciation, grammar).

Criteria are listed in the left column from most general to most specific moving down Table 4. The second column shows the variables used by ordinary (non-linguist) English speakers to assess L2 oral communication in EAP exams. The most important variables for these judges are shown in bold: *overall impression* and *content*, followed by two linguistic variables, *interaction* and *fluency* which depended on task type (paired speaker versus monologue). In the case of evaluation of veterinary students (column 3), the only study not to include an L2 dimension, both domain experts and applied linguists agreed on the pertinence of six additional criteria involving both linguistic and nonlinguistic dimensions, also indicated in boldface.[4] In physics (column 4), as noted, linguistic criteria were largely irrelevant to the discourse practices investigated, while medical experts (column 5) added two indigenous variables to existing linguistic test criteria. Finally research in civil aviation communication (column 6) questioned the relevance of two linguistic criteria (shown in italics) while underlying the shared nature of another (the burden of comprehension shared by speaker and hearer), and emphasising the importance of an indigenous variable, *professional knowledge*, alongside linguistic skills.

A number of variables which are common across LSP contexts and which cluster around the centre of the general-specific continuum are shown in the shaded area in the middle of Table 4. These include what language tests may simply label "interaction," but which the foregoing studies have fleshed out to include *discourse competence* (clarity, organisation of arguments, and transitions) and *interactional competence* (rapport, empathy, engagement, timing). In the following section I argue that it is here that further work on test development is both necessary and potentially fruitful, and here that a richer interpretation of communicative competence can help.

## Discussion

In the final section of this paper, I return to the notion of communicative competence in an attempt to synthesise the different positions described in the three domains of interest to the teaching and assessment of LSP: L2 research, L2 teaching, and LSP testing. From the foregoing discussion, it seems clear that questions of native-speaker norms and indigenous assessment criteria are of particular importance, motivating a return to a richer understanding of communicative competence. I examine each in turn before concluding with a proposal for a revised model of communicative competence to inform ongoing practice in LSP teaching and testing.

### L1 norms and L2 competence

One conclusion to be drawn from previous discussion is the extremely circumscribed focus of a good

---

[4] I include questioning skills as part of content for veterinary professionals rather than discourse structure since it seems to relate more closely to professional competence in this LSP context.

deal of work in both L2 research and L2 proficiency standards for teaching. As Widdowson (2017) points out, investigation of interlanguage development and CEFR proficiency scales has been largely confined to one small portion of Hymes' communicative competence: only knowledge, not ability for use (or indeed performance), only knowledge of "what is possible," not appropriateness, and only grammatical knowledge, not culture or communication. This narrow definition of knowledge is already apparent in the work of Canale and Swain (Table 2), where sociolinguistic competence is viewed as use, not knowledge. Similarly, although the goal of instructed L2 research is to study implicit learning made possible through engagement in communicative tasks (Long, 2017), communication is a means to an end, and the focus is on (cognitive) interlanguage development as opposed to interactional or intercultural competence.

In contrast, the sociocultural turn in L2 research (Firth & Wagner, 1997) rejects this view of L2 users as "defective communicators" and is supported by LSP research into indigenous communicative categories which shows that professional experts and ordinary language users pay little attention to formal accuracy when judging communicative success (Table 4). Content specialists (including average speakers who are not linguists) are generally able to agree on what constitutes effective communication in a specific domain and how to evaluate particular speakers. They generally do so with little reference to particular formal linguistic features, suggesting that they may be intuitively working with a broader model of communicative competence.

**Indigenous criteria as ability for use**

A comparison of the components of Hymes' (1972) model (Table 1) with indigenous assessment categories arising from recent LSP research (Table 4) suggests a number of useful intersections which justify a return to a broader and richer definition of communicative competence. Speakers who make a good general impression, show control of relevant content, and use effective discourse strategies to accommodate the needs of others in interaction may be considered to be demonstrating Hymes' knowledge of *what is appropriate*. Since these are all attested indigenous criteria, they also correspond to knowledge of *what is performed*. 'Timing' and 'accommodation strategies'[5], additional variables mentioned in medical, aviation, and scientific discourse, can be viewed as instances of *what is feasible*, since they correspond to a use of language to comply with physical constraints including safety issues and highly circumscribed discourse events.

Other indigenous categories in Table 4 pertain to Hymes' *ability for use*. The relevance of *affective factors* is attested in medical and veterinary interactions in terms such as rapport, empathy, tone of voice, and engagement. His *capacities in interaction* may be revealed in categories like 'overall impression' and 'demeanour', but also 'content accuracy' and 'admitting lack of knowledge', 'introductory/transitional/closing moves', and 'economy of expression', since all may depend on qualities of composure, presence of mind, and stage confidence which feature in Hymes' original model as aspects of ability for use. This interpretation of recent LSP work on indigenous assessment criteria goes some way to addressing earlier concerns with the operationalisation of this aspect of Hymes' model. Canale and Swain (1980) considered that the notion was insufficiently developed:

> We hesitate to incorporate the notion of ability for use into our definition of communicative competence for two main reasons: (i) to our knowledge this notion has not been pursued rigorously in any research on communicative competence (or considered directly relevant in such research), and (ii) we doubt that there is any theory of human action that can adequately

---

[5] Kim (2012) refers to accommodation strategies in the context of radiotelephonic communication between pilots and controllers where the main concern is precision and safety, not facework or politeness, hence the inclusion of this variable here.

explicate 'ability for use' and support principles of syllabus design intended to reflect this notion. (p. 7)

More recently, McNamara (1996) articulated a related concern:

Ability for use […] is more difficult to grasp, because we need to consider here a range of underlying language-relevant but not language-exclusive cognitive and affective factors (including general reasoning powers, emotional states and personality factors) which are involved in performance of communicative tasks. (p. 59)

On the basis of the research reviewed in the present paper, I suggest that the appropriate sphere of application for ability of use is the mid region of the continuum from the broadest definition of communicative competence (including global impressions, content knowledge and non-verbal behaviour) to the narrowest formal linguistic features (grammar and vocabulary). The central shaded area of Table 4, covering discourse strategies and interactional awareness, offers a promising locus for further investigation of this neglected dimension of communicative competence, perhaps in a model resembling Table 5.

**Table 5**: *Communicative Competence in LSP*

| Knowledge | | | Strategies for use | |
|---|---|---|---|---|
| linguistic | comprehension expression | grammar vocabulary | discourse | organisation argumentation delivery |
| | intelligibility | pronunciation fluency | | |
| pragmatic | appropriateness accommodation | | interaction | management engagement empathy |
| content | scientific occupational professional | | performance | demeanor non-verbal behaviour overall impression |

In Table 5 communicative competence is understood, following Hymes, as knowledge plus (strategies for) use, to use terms more in line with modern usage. Knowledge in LSP is of three main types: *linguistic*, covering traditional grammar, lexis, and intelligibility, *pragmatic*, reflecting sociolinguistic awareness related to rules of politeness and acceptability (following Widdowson, 2017), and *content* knowledge, corresponding to the professional or scientific knowledge of a specific content domain. Similarly, strategies for use are presented from the most circumscribed competences related to individual *discourse*, through *interactional* strategies to *performance* variables going beyond language.[6] These broad categories bear some comparison with Halliday's textual, ideational, and interpersonal dimensions (Halliday, 1978; Martin, 1992; Halliday, M. A. K., Matthiessen, & Halliday, M., 2014) and are indeed discussed in relation to communicative competence in the influential CLT paper by Breen and Candlin (1980). Since then Hallidayan lexicogrammar has had limited influence on the L2 teaching research reviewed here for reasons beyond the scope of this paper.

Concerning practical applications, LSP teaching might focus on the first two rows of Table 5 – linguistic and pragmatic knowledge with discourse and interactional strategies, while testing might be

---

[6] One anonymous reviewer suggests critical approaches in genre theory, new literacy studies and academic literacies (e.g., Lea & Street, 1998; Bhatia, 2008) have produced findings which are relevant here. I am inclined to side with Tribble (2017) who judges "there is little evidence in the research literature of their having had a significant impact on classroom teaching as yet" (p. 32).

more concerned with the bottom two – pragmatic and content knowledge plus interactional and performance strategies. The knowledge and strategies for use are relevant to both L1 and L2 speakers at all levels, since even linguistic competence may require attention for L1 speakers in professional lingua franca contexts. L1 users also need to learn intercultural skills and accommodation strategies for LSP communication. Harding (2014, p. 194-5) suggests that both L1 and L2 users need specific knowledge to allow accommodation (awareness of different varieties of English and new language patterns, including salient phonological features) as well as strategies for use such as negotiation of meaning, and noticing and repairing communication breakdowns.

In conclusion, this paper has argued for an expanded view of communicative competence which is more faithful to Hymes' (1972) original conception and reflects a number of advances in L2 research over the intervening five decades. One is the realisation that native-speaker norms are not the most relevant in LSP: formal linguistic accuracy is of little importance in any real-world context outside the language classroom. Another finding is that indigenous criteria for the assessment of communicative competence in both L1 and L2 contexts can produce categories of language use which offer reliable and valid indicators of speakers' performances. These criteria are comparable across disciplines and across discourse events and task types yet show little overlap with the linguistic criteria used in traditional EAP or LSP tests. This finding supports the view that our recent interpretations of communicative competence in language testing have failed to take the wider dimensions of this notion into account. I have argued that a middle ground between discrete formal linguistic criteria and broad extralinguistic factors is to be found in Hymes' original conceptions of both knowledge and ability for use, and that these dimensions can be usefully explored by researching interactional patterns and discursive practices in LSP communication.

## References

Bardovi-Harlig, K. (2000). *Aspect in second language acquisition: Form, meaning, and use*. Oxford: Blackwell Publishers.

Basturkmen, H., & C. Elder. (2004). The practice of LSP. In A. Davies & C. Elder (Eds.), *The handbook of applied linguistics*. Blackwell Publishing.

Bhatia, V. K. (2014). *Analysing genre: Language use in professional settings.* Routledge.

Bhatia, V. J. (2008). Towards critical genre analysis. In V. J. Bhatia, J. Flowerdew, & R. H. Jones (Eds.), *Advances in discourse studies*. Abingdon: Routledge.

Breen, M. P., & Candlin, C. N. (1980). The essentials of a communicative curriculum in language teaching. *Applied Linguistics*, *1* (2), 89-112.

Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics*, *1*, 1. doi:10.1093/applin/l.1.1.

Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

Corder, S. P. (1967). The significance of learner's errors. *International Review of Applied Linguistics in Language Teaching, 4*, 161-170. Retrieved from https://files.eric.ed.gov/fulltext/ED019903.pdf

Council of Europe. (2001). *Common European framework of reference for languages: Learning, teaching, assessment*. Strasbourg, France

Council of Europe (2005*). Reference level descriptions for national and regional languages*. Strasbourg, France.

Deygers, B., Van Gorp, K., & Demeester, T. (2018). The B2 level and the dream of a common standard. *Language Assessment Quarterly*, 1-15.

Douglas, D. (2001). Language for specific purposes assessment criteria: Where do they come from? *Language Testing, 18* (2), 171-185.

Douglas, D. (2013). ESP and assessment. In B. Paltridge. & S. Starfield (Eds.). *The handbook of*

*English for specific purposes*, (pp. 367-383).

Douglas, D., & Myers, R. (2000). Assessing the communication skills of veterinary students: Whose criteria? In *Fairness and validation in language assessment: Selected papers from the 19th Language Testing Research Colloquium*, Orlando, Florida (pp. 60-81). Cambridge: Cambridge University Press.

Dressen-Hammouda, D. (2012). Ethnographic approaches in ESP. In B. Paltridge, & S. Starfield (Eds.). *The handbook of English for specific purposes*. New York: Wiley.

Elder, C., McNamara, T., Kim, H., Pill, J., & Sato, T. (2017). Interrogating the construct of communicative competence in language assessment contexts: What the non-language specialist can tell us. *Language & Communication*, *57*, 14-21

Firth, A., & Wagner, J. (1997). On discourse, communication, and (some) fundamental concepts in SLA research. *The Modern Language Journal*, *81* (3), 285-300.

Fulcher, G. (2013). *Practical language testing*. Routledge.

Gregg, K. R. (1990). The variable competence model of second language acquisition, and why it isn't. *Applied Linguistics*, *11* (4), 364-383.

Gregg, K. R. (2010). Shallow draughts: Larsen-Freeman and Cameron on complexity. *Second Language Research*, *26* (4), 549-560.

Halliday, M. A. (1978). *Language as social semiotic*. London.

Halliday, M. A. K., Matthiessen, C., & Halliday, M. (2014). *An introduction to functional grammar*. Routledge.

Harding, L. (2014). Communicative language testing: Current issues and future research. *Language Assessment Quarterly*, *11* (2), 186-197.

Hawkins, J. A., & Filipović, L. (2012). Criterial features in L2 English. *Cambridge: CUP*.

Hulstijn, J. H. (2007). The shaky ground beneath the CEFR: Quantitative and qualitative dimensions of language proficiency. *The Modern Language Journal, 91*, 663–667. doi:10.1111/j.1540-4781.2007.00627_5.x

Hulstijn, J. H. (2014). The Common European Framework of Reference for Languages: A challenge for applied linguistics. *ITL-International Journal of Applied Linguistics*, *165* (1), pp. 3-18.

Hyland, K. (2002). Specificity revisited: How far should we go now? *English for Specific Purposes*, *21* (4), 385-395.

Hyland, K. (2006). *English for academic purposes: An advanced resource book*. Routledge.

Hymes, D. (1972). On communicative competence. In J.B. Pride and J. Holmes (Eds.). *Sociolinguistics*. Harmondsworth: Penguin.

Jacoby, S. W. (1999). *Science as performance: Socializing scientific discourse through the conference talk rehearsal*. (Unpublish doctoral dissertation). University of California, Los Angeles.

Jacoby, S., & McNamara, T. (1999). Locating competence. *English for Specific Purposes, 18* (3), 213-241.

Johns, A. M. (2013). The history of English for specific purposes research. In B. Paltridg & S. Starfield (Eds.). *The handbook of English for specific purposes* (pp. 5-30). Wiley.

Kim, H. (2012). *Exploring the construct of aviation communication: A critique of the ICAO language proficiency policy.* (Unpublished doctoral thesis). University of Melbourne.

Kim, H., & Billington, R. (2016). Pronunciation and comprehension in English as a lingua franca communication: Effect of L1 influence in international aviation communication. *Applied Linguistics*, *2* (1), doi.org/10.1093/applin/amv075

Kramsch, C. (2000). Second language acquisition, applied linguistics, and the teaching of foreign languages. *The Modern Language Journal*, *84* (3), 311-326.

Lakoff, G., & Johnson, M. (1999). *Philosophy in the Flesh*. New York: Basic Books.

Lantolf, J. P. (Ed.). (2000). *Sociocultural theory and second language learning*. Oxford: Oxford University Press.

LarsenFreeman, D., & Cameron, L. (2008). *Complex systems and applied linguistics*. Oxford: Oxford

University Press.

Lea, M. R., & Street, B. V. (1998). Student writing in higher education: An academic literacies approach. *Studies in Higher Education*, *23* (2), 157-172.

Lockwood, J. (2012). Developing an English for specific purpose curriculum for Asian call centres: How theory can inform practice. *English for Specific Purposes*, *31* (1), 14-24.

Long, M. (2017). Instructed second language acquisition (ISLA): Geopolitics, methodological issues, and some major research questions. *ISLA*, *1* (1), 7-44.

Martin, J. R. (1992). *English text: System and structure*. John Benjamins Publishing.

McNamara, T. F. (1996). *Measuring second language performance*. Addison Wesley Longman.

McLelland, N., & Smith, R. (2014). Introduction: Building the history of language learning and teaching (HoLLT). *Language and History, 57*, 1-9.

O'Keeffe, A., & Mark, G. (2017). The English grammar profile of learner competence. *International Journal of Corpus Linguistics, 22* (4), 457-489.

Paltridge, B. (2012). Genre and English for specific purposes. In B. Paltridge, & S. Starfield (Eds.). *The handbook of English for specific purposes*. New York: Wiley.

Pill, T.J.H. (2013). What doctors value in consultations and the implications for specific-purpose language testing (Unpublished doctoral thesis). University of Melbourne, Australia.

Sato, T. (2014). Linguistic laypersons' perspective on second language oral communication ability Doctoral dissertation, University of Melbourne.

Selinker, L. (1972). Interlanguage. *International Review of Applied Linguistics in Language Teaching*, *10* (1-4), 209-232.

Selinker, L., & Douglas, D. (1985). Wrestling with 'context' in interlanguage theory. *Applied Linguistics*, *6* (2), 190-204.

Skehan, P. (2009). Modelling second language performance: Integrating complexity, accuracy, fluency, and lexis. *Applied linguistics*, *30* (4), 510-532.

Spada, N. (2015). SLA research and L2 pedagogy: Misapplications and questions of relevance. *Language Teaching*, *48* (1), 69-81.

Smith, R., & McLelland, N. (2018). Histories of language learning and teaching in Europe. *The Language Learning Journal*, *46* (1), 1-5.

Tarone, E. (1983). On the variability of interlanguage systems. *Applied Linguistics*, *4* (2), 142-164.

Thewissen, J. (2013). Capturing L2 accuracy developmental patterns: Insights from an error-tagged EFL learner corpus. *The Modern Language Journal*, *97* (S1), 77-101.

Tribble, C. (2017). ELFA vs. Genre: A new paradigm war in EAP writing instruction? *Journal of English for Academic Purposes*, *25*, 30-44.

Van Ek, J., & Alexander, L. G. (1975). Threshold level English in a European unit/credit system for modern language learning by adults. Strasbourg: Council of Europe.

Widdowson, H. (2017). Disciplinarily and disparity in applied linguistics. British Association for Applied Linguistics Annual Conference, August. https://youtu.be/choufPZm1O8 (31:47)

Wilkins, D. A. (1973). The linguistic and situational content of the common core in a unit/credit system. *Systems development in adult language learning*. Strasbourg: Council of Europe. Retrieved from https://files.eric.ed.gov/fulltext/ED136550.pdf.

**Author Biodata**

**Shona Whyte** is Professor of English at Université Côte d'Azur, CNRS, BCL, France. Her research interests within second language studies involve English for specific purposes, computer-assisted language learning, and language teacher education. Recent books include *New Developments in ESP Teaching and Learning Research (*with C Sarré, Researchpublishing.net, 2017)*, Implementing and*

*Researching Technological Innovation in Language Teaching* (Palgrave, 2015), and *Teaching Languages with Technology* (with E. Cutrim Schmid, Bloomsbury, 2014).