

From Linguistic Skills to Pragmatic Competence: The Role of Functional Adequacy in Task-Based Teaching and Learning

Gavin BUI

The Hang Seng University of Hong Kong, HKSAR

Chi Him WONG

The Hang Seng University of Hong Kong, HKSAR

Despite advancements in the tripartite framework of task-based second language performance assessment, namely complexity, accuracy and fluency (CAF), functional adequacy (FA) has not gained sufficient attention in the field. It can be argued that language learners may be able to produce linguistically advanced speech that may not be pragmatically appropriate or achieve the goals of the task they set out to complete. In light of this, the current paper explains the notion of functional adequacy, stresses its importance in being integrated into the conventional CAF language assessment parameters, and outlines the relationship between CAF and FA. Finally, it proposes areas for further research and suggestions for the use of FA in task-based second language teaching.

Introduction

“Colorless green ideas sleep furiously” is a well-known example in Chomsky’s (1957) seminal book *Syntactic Structures* that exemplifies a case of the co-existence of grammatical correctness and semantical emptiness. Neuroscientists have long noticed Wernicke’s aphasia patients who demonstrate fluent, effortless speech characterized by normal tempo and grammatical structures, but that is difficult to understand; the speech typically lacks content or meaning and thus becomes nonsensical to hearers. It appears that the popular belief among second language (L2) teachers and researchers that syntactic abilities ought to be prioritized in language teaching is not watertight; the L2 teaching profession should seek to develop learners’ communicative competence, both linguistically and socio-linguistically. Unfortunately, this issue has not been well addressed in past L2 teaching and learning literature; it has been largely neglected even in the field of task-based language teaching (TBLT) which, since its first proposal, has aimed to develop real-world communicative abilities. This article discusses the conventional constructs of complexity, Accuracy and Fluency (CAF) (Bui & Skehan, 2018; Housen & Kuiken, 2009; Michel, 2017) and their inadequacies, and argues why Functional Adequacy (FA) should merit a place in task performance assessment (Kuiken & Vedder, 2018; Pollotti, 2009). It will then examine various task features of FA performance, followed by the relationship between FA and CAF measures. The article will then discuss gaps for future research. Finally, it concludes with some pedagogical implications for L2 teaching.

Conventional Task Performance Assessment: CAF and CALF

Learners’ L2 task performance has been typically measured along the dimensions of complexity, accuracy and fluency (CAF) as was proposed by Skehan (1996) and further elaborated by Housen and Kuiken (2009), Michel (2017), and Bui and Skehan (2018). Skehan (2009), among others, further called for the inclusion of lexical aspects in the CAF framework, which then became CALF (Bui, 2021). The CAF and its recent version CALF have greatly influenced many TBLT researchers, providing them with a set of widely recognized indices that will be further discussed below.

As an area of learners' L2 performance, complexity was defined by Housen and Kuiken (2009) as the size, elaborateness, richness and diversity of the L2 sample that an L2 learner produces. Greater language complexity can be a sign of L2 interlanguage development (Norris & Ortega, 2009), but it can also indicate learners' inclination to experiment with new linguistic structures (Bui & Skehan, 2018). Michel (2017) distinguished between three types of complexity: developmental, cognitive and linguistic complexity. Among them, linguistic complexity is the dimension subject to empirical measurement, hence its common use in task-based performance assessment. Unfortunately, valid and reliable indices for linguistic complexity have not always been agreed upon, but the task literature has seemed to come to terms with two broad categories: grammatical complexity and lexical complexity (Bulté & Housen, 2012). The literature has seen grammatical complexity typically measured as the length of a unit, such as a clause or an AS unit (the Analysis of Speech Unit proposed by Foster et al., 2000), or as the ratio of subordinate clauses to all clauses (Norris & Ortega, 2009).

On the other hand, lexical complexity is usually assessed in terms of lexical diversity, lexical sophistication and lexical density in task research (Bui, 2021; Bulté & Housen 2012), with various indices. Lexical diversity reflects the extent to which speakers draw on a larger vocabulary and avoid recycling the same set of words. To avoid the influence of text length on the raw type-token ratio (TTR), lexical diversity is often measured through various corrected TTRs, such as *VocD* (Malvern & Richards, 2002) or *MTLD* (Measure of Textual Lexical Diversity) (McCarthy, 2005). The second dimension of lexical complexity, lexical sophistication, is indexed by a range of word frequency-based measures (Read, 2000) to assess the degree to which learners use less frequent words (Skehan, 2009). Common measures include the Lexical Frequency Profile (LFP, Laufer & Nation, 1995), which outputs a range of frequency bands, and the *Lambda*, which is a single value calculated through *P_lex* (Meara & Bell, 2001). The third lexical complexity aspect is lexical density, commonly measured through the ratio of content words to all words produced by a learner, to indicate "the density of information in any passage of text, according to how tightly the lexical items (content words) have been packed into the grammatical structure" (Halliday & Martine, p. 76).

Accuracy is the extent to which a learner's performance conforms to the grammatical system of a second language that is being learned.

Greater accuracy as shown in L2 production may indicate development in syntactic proficiency, but it could also suggest a conservative approach to language learning, such as the avoidance strategy of speaking in shorter utterances. Accuracy can be measured based on either specific error counts (e.g., errors in past tense) or general measures (e.g., the ratio of error-free clauses). More recently, different attempts (Foster & Wigglesworth, 2016; Kuiken & Vedder, 2008) have been made to provide accuracy metrics with weightings on different types of errors. Prior research (e.g., Bui & Skehan, 2018) has argued for the inclusion of different accuracy indices to assess: (1) the number of errors, (2) their severity, and (3) the relationship of the errors to clausal length to cover more possibilities in accuracy performance.

The last area in CALF, fluency, refers to the ease, eloquence and smoothness of speech or writing. Segalowitz (2010) distinguished between three types of fluency, namely, cognitive fluency, utterance fluency and perceived fluency. Utterance fluency is measurable and thus becomes the most relevant to task performance. Tavakoli and Skehan (2005) identified three aspects of L2 utterance fluency: (1) speed, e.g., the number of words/syllables per minute; (2) breakdown, e.g., number, length and location of pauses; and (3) repair, e.g., false starts, reformulation, repetitions and replacement. A more fine-grained categorization of fluency was presented in Bui and Huang (2018), in which they proposed to measure L2 fluency in terms of speed, stretch, voicing, mid-clause unfilled pauses, end-of-clause (dependent and independent clauses) unfilled pauses, filled pauses and repairs. They argued that in addition to the frequency and length of pauses, the position of pausing matters significantly. Mid-clause pausing was found to be a feature of L2 speaking while pausing at clausal boundaries would sound more natural and nativelike. In general, a focus on fluency suggests attention to meaning expression rather than forms in real-time communication.

Over the past 30 years, research on complexity, accuracy, lexis and fluency (CALF) has yielded fruitful results, with a set of measures emerging to be widely accepted in TBLT. However, a closer examination of CALF reveals a major research lacuna in that functional adequacy, or communicative effectiveness, has not been included in the exploration of learner performance (Pollatti, 2009, Kuiken & Vedder, 2018). We will now turn to the discussion of FA as an indispensable area of task performance.

More Comprehensive Assessment: From CALF to CALFFA

In the current literature, advanced performance is often equated with the learners' abilities to speak or write fluently and accurately, and with complex syntactic structures. However, as discussed above, accurate command of the linguistic aspects of the language alone does not necessarily entail success in real-world communication; one can always speak fluent and grammatically complex and correct sentences, yet without achieving any communicative goals. Pallotti (2009) stated that psycholinguistic factors, such as automaticity of cognitive efficiency and memory, may not necessarily be the causes of variations in CALF. He argued that the task's semantic and pragmatic demands could instead be the potential reason behind the variations observed. For instance, a learner may use a simple sentence because of their lack of proficiency, but it may also be a deliberate decision given the setting of the task, as determined by the speaker, for instance, when they are speaking to a toddler; by contrast, complex sentences may not always be appropriate in day-to-day spoken communication. Functional adequacy, defined as "the degree to which a learners' performance is more or less successful in achieving the task's goals efficiently" (Pallotti, 2009, p. 596), is thus a crucial factor in assessing task performance, but one which has unfortunately been overlooked in most existing TBLT literature to date (Kuiken & Vedder, 2018). It appears that Pallotti's (2009) conceptual endeavor for functional adequacy can complement the conventional CALF, which helps to establish a more comprehensive assessment framework for task-based language performance. The proposal of adding FA into the CALF measures is also in line with Leung's (2005) call for a reorientation for communicative competence. A move from CALF to CALFFA conforms to the notion of language proficiency as promulgated in the Common European Framework of Reference (CEFR) for languages. The CEFR has defined language proficiency not only in terms of functions, domains and roles of language users can handle (the "what") but also in terms of the quality of language proficiency (the "how well") (Hulstijn, 2007).

There is, then, a research gap regarding the link between the communicative adequacy of L2 production and the linguistic forms of the conveyed message. One of the reasons for this gap is possibly due to the lack of a precise and systematic definition of communicative

adequacy as a construct (Kuiken, Vedder & Gilabert, 2010). For example, Hymes (1966) had proposed in one of his early works the notion of “communicative competence,” contrasted with the Chomskyan “linguistic competence.” In more recent discussions, a wide range of different definitions have been brought up, such as intercultural competence (Hismanoglu, 2011), communicative effectiveness (Sato, 2012), communicative adequacy (Kuiken et al. 2010), communicative functionality (Fragai, 2003), “communicative competence” (Leung 2005), “pragmatic proficiency” (Takimoto 2020), “functional adequacy” (De Jong 2012) and “communicative adequacy” (Kuiken et al. 2010). While such diversity may vitalize research, it is nevertheless an impediment to arriving at a standardized approach to the study of functional adequacy for a common purpose.

Current Definitions of Functional Adequacy

As discussed in the last section, numerous terms have been used in the literature in reference to what may be considered similar concepts in communicative competence or adequacy. All of these terms appear to have the same underlying notion of “functional adequacy” (FA), a term that has been gaining traction as an umbrella term for the aforementioned subject matters. However, as mentioned above, there is currently no consensus about its precise definition. Definitions given include “knowledge and employment of both linguistic and interactional resources in social contexts” (Révész et al., 2016, p. 829) and “the degree to which a learners’ performance is more or less successful in achieving the task’s goals efficiently” (Pallotti, 2009, p. 596). Kuiken and Vedder (2018) have quoted definitions like “successful information transfer” (Upshur & Turner, 1995), pragmatic appropriateness (McNamara & Roever, 2007) and text coherence and cohesion (Knoch, 2009), while their favored definition has been given as “successful task completion” (Kuiken & Vedder, 2018, p. 268).

Less broadly, specific definitions distinctly with regards to spoken or written FA could also be found in previous research. In Takimoto’s (2020) work on spoken FA, he has given his definition as “understanding the speaker’s intended meaning in a particular context” (p. 1). As for defining written FA, Kuiken and Vedder’s definitions have been referenced by numerous authors, with Herraiz-Martínez (2018) employing the definition from their 2016 study of being an “interpersonal construct

which measures L2 writing in terms of successful task fulfilment” (Kuiken & Vedder, 2016, as cited in Herraiz-Martínez, 2018, p. 19), while Nuzzo and Bove (2020) cited their definition from the 2017 paper, which defines it as a “successful task completion of A in conveying a message to B and in relation to the conversational maxims of Grice” (Kuiken & Vedder, 2017, as cited in Nuzzo & Bove, 2020, p. 10).

Despite the variations among the definitions given above, one can find that there is a common notion of success in conveying information (Kuiken and Vedder, 2017; Upshur & Turner, 1995) or otherwise the extent to which a task has been successfully completed (Révész et al., 2016; Pallotti, 2009; Kuiken & Vedder, 2016). In light of such commonality identified in prior studies, Kuiken and Vedder (2018, pp. 282–283) proposed the following parameters in assessing FA in task performance, using a rating scale from one to six for each:

1. “Content: Is the number of information units provided in the text adequate and relevant?”
2. “Task Requirements: Have the task requirements been fulfilled successfully (e.g. genre, speech acts, register)?”
3. “Comprehensibility: How much effort is required to understand text purpose and ideas?”
4. “Coherence and cohesion: Is the text coherent and cohesive (e.g. cohesive devices, strategies)?”

They found that this approach to the assessment of functional adequacy had been reliable and valid from application to the writing of L2 Dutch and L2 Italian learners in the Netherlands. Unmistakably, such works by Pallotti (2009), Révész et al. (2016) and Kuiken and Vedder (2018) are commendable pioneers on these topics, which have laid important foundations for later research into this area.

A shift of attention from linguistic to FA performance in tasks is associated with growing interest in interlanguage pragmatics instruction and assessment (Taguchi & Kim, 2018). Pragmatics concerns how language is used in context and how meaning is conveyed appropriately. As pointed out by Youn (2013), a common theme underlying different approaches to pragmatics is that speakers could be regarded as “individual rational actors who choose their means to meet the actor’s intended goals” following encoding/decoding certain communicative protocols (p. 9). It transpires that intentionality of the speaker and recognition of such intentionality by the listener constitute the core of interactional

pragmatics, which in turn lies at the heart of functional adequacy in task performance as it aims to develop real-world communicative skills. Its importance in TBLT has been reflected in recent publications on the use of tasks in developing L2 pragmatic skills (e.g., Gilabert & Barón, 2018; Révész et al., 2016; and the whole edited volume by Taguchi & Kim (2018) on teaching and testing pragmatics in TBLT). FA should therefore be defined as a communicative-goal-driven, context-sensitive, and culture-specific construct. It appears that current definitions of FA have met the first two criteria but a more comprehensive definition that involves cross-cultural awareness is still lacking; the same seems also true for the existing measurement of FA.

Relationships Between FA and CAF

Spoken FA

One of the few studies that have investigated the relationship between linguistic forms and spoken FA would be the study by Révész et al. (2016). This research has explored in great detail the degree to which the linguistic forms of CAF predict speaking functional adequacy, and whether or not a task type and language proficiency have an effect on mediating these relationships. The study involved 20 native speakers and 80 L2 users, who were placed into four groups based on their proficiency levels. After collecting speech samples from five distinct task types, the samples were assessed in terms of CAF and spoken FA. Their findings suggested that false starts for more proficient speakers and the frequency of filled pauses (an aspect of fluency) were the strongest predictors of FA. Other CAF measures that predict FA with less reliability included subscales of all three CAF features: complexity (lexical diversity, overall syntactic complexity, subordination ratio and frequency of conjoined clauses), accuracy (both overall accuracy and accuracy of connectors) and fluency (unfilled pauses and speech rate). Vasylets et al.'s (2020) study also had similar findings in this area, finding spoken FA to have a consistent relationship with lexical complexity, fluency and accuracy. Additionally, Révész et al.'s research explored the extent to which FA was affected by task types, by administering tasks that were grouped respectively into complaint, refusal, narrative, advice, and summary, and the results suggested that the relationship between the CAF measures

and FA remains consistent regardless of task types. Thus, the study had not found task types to have any significant impact on the relationship between linguistic features and adequacy.

However, in contrast to the findings from Révész et al., De Jong et al.'s studies (2007; 2012a) suggested that complex tasks result in lower oral FA for L2 speakers. This discrepancy may be attributed to the fact that Révész et al. did not group task types by task complexity. Nonetheless, under Levelt's model (Levelt 1989; Levelt, Roelofs & Meyer 1999, as cited in De Jong et al., 2012a), when speakers struggle with fluency, they use most of their mental resources on that aspect, and as such do not have enough attentional resources to maintain FA, suggesting that complex tasks negatively impact fluency, which in turn hinders FA.

The research by Pallotti (2019) also suggests that tasks requiring a higher linguistic complexity lead to increased task difficulty. This is also echoed in De Jong et al.'s studies (2007; 2012a), where it could be inferred that complex language needs to be used in more complex tasks, and that complex tasks negatively affect the FA of first and second language users alike, again suggesting that complexity indirectly impacts FA.

Written FA

There have been fewer studies exploring written FA in task-based L2 teaching literature, as noted by Herraiz-Martínez (2018). Many of the studies referenced have not employed the CAF rating system, borrowing instead from one of the revisions of Kuiken and Vedder's scale. Nonetheless, both Kuiken et al.'s (2010) and Herraiz-Martínez's (2018) studies found that high proficiency is a necessity for high written FA.

From Kuiken (n.d.), overall ratings of learners' linguistic complexity performances show significant similarities to their written FA ratings and are correlated with lexical diversity and accuracy, though not with syntactic complexity. This suggests an intimate relationship between linguistic performance and FA, specifically in the lexical diversity and accuracy aspects.

Vasylets et al.'s (2020) study, like the case with spoken FA, linked written FA with lexical complexity and fluency, in addition to the number of ideas expressed, which was termed "propositional complexity" in the paper. The study, however, finds that this relationship remains consistent regardless of task complexity.

Gaps for the Future

Despite some growing interest in FA in task-based language teaching, there still exist many areas that are worth further exploring. First, there is a need for general measures for rating the functional adequacy of L2 performance (Kuiken & Vedder, 2017). There are two areas of interest for this: the lack of a clear definition of FA (the validity issue) and the difficulty in obtaining consistent human ratings (the reliability issue). Second, there is a paucity of research studying the links between FA and CALF; the amount and types of linguistic features that signify communicatively adequate speech continues to remain mostly unresearched. As Révész et al. (2016, p. 829) mentioned:

Despite the importance attributed to tasks as promoters and assessments of communicative adequacy, the bulk of task-related SLA research has been directed at examining the linguistic outcomes of task performance, expressed in terms of syntactic and lexical complexity, accuracy, and fluency, without considering how these features may relate to communicative adequacy.

Thirdly, the FA framework proposed by Vedder and Kuiken (2017) may not be applicable to creative writing tasks, especially the “content” and “task fulfilment” dimensions. One aspect of Kuiken and Vedder’s measurement for task fulfilment is rating participants’ output for the degree in which the task requirements were reached. However, task requirement has a presupposed goal in which participants are expected to attain, or work towards attaining. Such is true of tasks such as responding to a tourist who is asking for directions, where task fulfilment would be rated based on how much of the required information the participant has successfully relayed to the tourist. However, as a predetermined goal does not exist for creative writing, it becomes challenging to measure the degree to which participants’ outputs reach the goals of the task.

Regarding the “content” dimension of Kuiken and Vedder’s FA measurement, the dimension is judged according to the adequacy and consistency of the presented ideas. However, as no model framework exists for how a piece of creative writing should be, there are no grounds for how idea adequacy is to be rated. Attempting to rate content based on the quality of the work, as a means of measuring the adequacy and sufficiency of ideas or how successful they are at “being creative,” would leave the results susceptible to rating subjectivity. Creative writing tasks, by definition, require a high degree of improvisation and

unpredictability. A set of pre-set content and task goals may run counter to the purpose of creative writing “task fulfilment.” It seems that a contextualized and specialized set of FA rating scales would be necessary for tasks of this nature.

Another important topic for further studies would be cross-cultural awareness. L2 speakers of different cultural backgrounds may see FA differently, as do FA raters. However, the focus of most existing frameworks of FA assessment appears to be on the “task” *per se*, or rather, the micro-aspects of performance within the European framework; the macro-aspects, or the bigger context of cross-cultural and cross-linguistic awareness, do not appear to have received much attention in the proposed models.

Pedagogical Implications

As discussed above, FA measurements could prove to be a useful tool in the analysis of speaker proficiency. Kuiken and Vedder (2018) tested their rating methods with non-expert raters and found that the raters had high reliability among each other, and thus deemed their scale to be a step towards a reliable tool for measuring FA in L2 speakers. Thus, some form of Kuiken and Vedder’s scale may help to better assess learners’ language competence. In addition, Kuiken and Vedder (2019) have added that FA measurement results can be used as a diagnostic tool for giving dedicated feedback to students, though they state that there is still a need for methodology standardisation, rater training, and data analysis of FA.

Herraiz-Martínez’s (2018) study investigated the effect of teaching strategies on FA and found that group discussion activities and peer-reviewing contribute to proficiency, and in turn, FA. Her study also found that the intensity of EMI (English medium instruction) also positively affected FA. Takimoto’s (2020) study concluded that societal appropriateness strategies taught using pictures and metaphors resulted in higher FA than when taught in a rote learning method. Ryan and Granville (2020), however, warn against the use of film dialogue as a model for teaching social appropriateness to students, as they had found that turn sequencing in films differs significantly from real-world communicative norms. Thus, social appropriateness may be taught in the classroom with more picture-based and student-to-student interactive methods to improve learners’ FA.

Leung (2005) states that there should be a better interpretation of communicative competence, taking into account the cultural and societal norms of native and non-native speakers of English. Particularly, as there is a high level of variation between different forms of the same language, sticking to a single idealized perceived native correctness for both grammar and social norms would be problematic. Leung proposes more focus on real-world communication in terms of a higher focus on real-life social appropriateness, and on accommodating communicative aims and non-standard varieties alongside the standard, in English language education, rather than solely grossly generalized social appropriateness and language based on a single idealized form.

Conclusion

This paper discusses the importance of Functional Adequacy in task-based language teaching when assessing learners' L2 proficiency in addition to the conventional CALF paradigm. A speaker exhibiting simpler complexity may be a more socially appropriate choice for a given setting and would show a greater, rather than smaller, command of the use of the language by a speaker when done correctly. In contrast, a learner can produce grammatically accurate and fluent speeches without being pragmatically appropriate or conveying any meaningful content. These demonstrate the usefulness of including FA into task performance or even general proficiency measurements. The rating method proposed by Kuiken and Vedder could be employed in classrooms as a reliable measurement for FA. Current findings suggest that the CAF parameters tend to share a close relationship with FA, with higher CAF performances tending to be seen also with higher FA performance. Task complexity appears to inversely impact FA, with more complex tasks resulting in lower FA by L2 users. A shift in focus from the linguistic accuracy of a single form to adhering to real-world appropriateness may be advisable for the future of English language education.

Acknowledgements

This research was supported by generous funding (Ref. No. UGC/FDS14/H13/20) from the Research Grants Council, University Grants Committee, Hong Kong.

References

- Bui, G. (2021). Influence of learners' prior knowledge, L2 proficiency and pre-task planning time on L2 lexical complexity. *IRAL-International Review of Applied Linguistics in Language Teaching*, 59(4), 543–567. <https://doi.org/10.1515/iral-2018-0244>
- Bui, G., & Huang, Z. (2018). L2 fluency as influenced by content familiarity and planning: Performance, methodology and pedagogy. *Language Teaching Research*, 22(1), 94–114. <https://doi.org/10.1177/1362168816656650>
- Bui, G., & Skehan, P. (2018). Complexity, fluency, and accuracy. In J. Liantas (Ed.), *The TESOL encyclopedia of English language teaching* (pp. 1–8). John Wiley & Sons, Inc. <https://doi.org/10.1002/9781118784235.eelt0046>
- Bulté, B., & Housen, A. (2012). Defining and operationalising L2 complexity. In A. Housen, F. Kuiken, & I. Vedder (Eds.), *Dimensions of L2 performance and proficiency: Investigating complexity, accuracy and fluency in SLA* (pp. 21–46). John Benjamins. <https://doi.org/10.1075/llt.32.02bul>
- Chomsky, N. (1957). *Syntactic structure*. Mouton.
- De Jong, N. H., Steinel, M., Florijn, A. F., Schoonen, R., & Hulstijn, J. H. (2007). The effect of task complexity on fluency and functional adequacy of speaking performance. *Journal of Educational Psychology*, 99(3), 477–491.
- . (2012a). Facets of speaking proficiency. *Studies in Second Language Acquisition*, 34(1), 5–34. <https://doi.org/10.1017/S0272263111000489>
- . (2012b). The effect of task complexity on functional adequacy, fluency and lexical diversity in speaking performances of native and non-native speakers. In A. Housen, F. Kuiken, & I. Vedder (Eds.), *Dimensions of L2 performance and proficiency: Complexity, accuracy and fluency in SLA* (pp. 121–142). John Benjamins. <https://doi.org/10.1075/llt.32.06jon>
- Foster, P., Tonkyn, A., & Wigglesworth, G. (2000). Measuring spoken language: A unit for all reasons. *Applied Linguistics*, 21, 354–375. <https://doi.org/10.1093/applin/21.3.354>
- Foster, P., & Wigglesworth, G. (2016). Capturing accuracy in second language performance: The case for a weighted clause ratio. *Annual Review of Applied Linguistics*, 36, 98–116. <https://doi.org/10.1017/S0267190515000082>
- Gilabert, R., & Barón, J. (2018). Independently measuring cognitive complexity in task design for interlanguage pragmatics development. In N. Taguchi & Y. Kim (Eds.), *Task-based approaches to teaching and assessing pragmatics* (pp. 160–190). John Benjamins.
- Halliday, M. A. K., & Martine, J. R. (1993). *Writing science: Literacy and discursive power*. University of Pittsburgh Press.
- Herreraiz-Martínez, A. (2018). *Functional adequacy: The influence of English-medium instruction, English proficiency and previous language learning experience* [Unpublished doctoral dissertation, University Jaume I]. <http://dx.doi.org/10.6035/14110.2018.368282>

- Hismanoglu, M. (2011). An investigation of ELT students' intercultural communicative competence in relation to linguistic proficiency, overseas experience and formal instruction. *International Journal of Intercultural Relations*, 35(6), 805–817. <https://doi.org/10.1016/j.ijintrel.2011.09.001>
- Housen, A., & Kuiken, F. (2009). Complexity, accuracy and fluency in second language acquisition. *Applied Linguistics* 30(4), 461–473. <https://doi.org/10.1093/applin/amp048>
- Hulstijn, J. H. (2007). The shaky ground beneath the CEFR: Quantitative and qualitative dimensions of language proficiency. *The Modern Language Journal*, 91(4), 663–667. https://doi.org/10.1111/j.1540-4781.2007.00627_5.x
- Kuiken, F. (n.d.). The assessment of functional adequacy and the relation with linguistic complexity in L2 writing. https://www.arts.kuleuven.be/ling/taalenonderwijs/valilex/documenten/kuiken_150223
- Kuiken, F., & Vedder, I. (2008). Cognitive task complexity and written output in Italian and French as a foreign language. *Journal of Second Language Writing*, 17(1), 48–60. <https://doi.org/10.1016/j.jslw.2007.08.003>
- . (2017). Functional adequacy in L2 writing: Towards a new rating scale. *Language Testing*, 34(3), 321–336. <https://doi.org/10.1177%2F0265532216663991>
- . (2018). Assessing functional adequacy of L2 performance in a task-based approach. In N. Taguchi & Y. Kim (Eds.), *Task-based approaches to teaching and assessing pragmatics* (pp. 266–285). John Benjamins. <https://doi.org/10.1075/tblt.10.11kui>
- . (2019). *From CAF to CAFFA. Measuring linguistic performance and functional adequacy in task-based language teaching*. Keynote speech presented at Eighth International Conference on Task-Based Language Teaching (August 19–21), Carleton University, Ottawa, Canada.
- Kuiken, F., Vedder, I., & Gilabert, R. (2010). Communicative adequacy and linguistic complexity in L2 writing. In I. Bartning, M. Martin, & I. Vedder (Eds.), *Communicative proficiency and linguistic development: Intersections between SLA and language testing research* (pp. 81–100). European Second Language Association.
- Laufer, B., & Nation, P. (1995). Vocabulary size and use: Lexical richness in L2 written production. *Applied Linguistics*, 16, 307–322. <https://doi.org/10.1093/applin/16.3.307>
- Leung, C. (2005). Convivial communication: Recontextualizing communicative competence. *International Journal of Applied Linguistics*, 15(2), 119–144. <https://doi.org/10.1111/j.1473-4192.2005.00084.x>
- Malvern, D., & Richards, B. (2002). Investigating accommodation in language proficiency interviews using a new measure of lexical diversity. *Language Testing*, 19(1), 85–104. <https://doi.org/10.1191%2F0265532202lt221oa>

- McCarthy P. M. (2005). *An assessment of the range and usefulness of lexical diversity measures and the potential of the measure of textual, lexical diversity* [PhD dissertation, The University of Memphis]. ProQuest Dissertations Publishing (UMI No. 3199485). [https://doi.org/10.1044/1058-0360\(2013/12-0083\)](https://doi.org/10.1044/1058-0360(2013/12-0083))
- Meara, P., & Bell, H. (2001). P_Lex: A simple and effective way of describing the lexical characteristics of short L2 texts. *Prospect*, 16(3), 5–19.
- Michel, M. (2017). Complexity, accuracy and fluency in L2 production. In S. Loewen & M. Sato (Eds.), *The Routledge handbook of instructed second language acquisition* (pp. 50–68). Routledge. <https://doi.org/10.4324/9781315676968>
- Norris, J. M., & Ortega, L. (2009). Towards an organic approach to investigating CAF in instructed SLA: The case of complexity. *Applied Linguistics*, 30, 555–578. <https://doi.org/10.1093/applin/amp044>
- Nuzzo, E., & Bove, G. (2020). Assessing functional adequacy across tasks: A comparison of learners' and native speakers' written texts. *EuroAmerican Journal of Applied Linguistics and Languages*, 7(2), 9–27. <http://doi.org/10.21283/2376905X.12.175>
- Pallotti, G. (2009). CAF: Defining, refining and differentiating constructs. *Applied Linguistics*, 30(4), 590–601. <https://doi.org/10.1093/applin/amp045>
- . (2019). An approach to assessing the linguistic difficulty of tasks. *Journal of the European Second Language Association*, 3(1), 58–70. <https://doi.org/10.22599/jesla.61>
- Read, J. (2000). *Assessing vocabulary*. Cambridge University Press.
- Révész, A., Ekiert, M., & Torgersen, E. N. (2016). The effects of complexity, accuracy, and fluency on communicative adequacy in oral task performance. *Applied Linguistics*, 37(6), 828–848. <https://doi.org/10.1093/applin/amu069>
- Ryan, J., & Granville, S. (2020). The suitability of film for modelling the pragmatics of interaction: Exploring authenticity. *System*, 89, 1–13. <https://doi.org/10.1016/j.system.2019.102186>
- Segalowitz, N. (2010). *Cognitive bases of second language fluency*. Routledge. <https://doi.org/10.4324/9780203851357>
- Tavakoli, P., & Skehan, P. (2005). Strategic planning, task structure, and performance testing. In R. Ellis (Ed.), *Planning and task performance in a second language* (pp. 239–277). John Benjamins. <https://doi.org/10.1075/llt.11.15tav>
- Taguchi, N., & Kim, Y. (Eds.). (2018). *Task-based approaches to teaching and assessing pragmatics*. John Benjamins.
- Skehan, P. (2009). Lexical performance by native and non-native speakers on language-learning tasks. In B. Richards, H. Daller, D. D. Malvern, & P. Meara (Eds.), *Vocabulary studies in first and second language acquisition: The interface between theory and application* (pp. 107–124). Palgrave Macmillan. https://doi.org/10.1057/9780230242258_7

- Takimoto, M. (2020). Investigating the effects of cognitive linguistic approach in developing EFL learners' pragmatic proficiency. *System*, 89, 1–14. <https://doi.org/10.1016/j.system.2020.102213>
- Vasylets, O., Gilabert, R., & Manchón, R. M. (2020). Task modality, communicative adequacy and CAF measures: The moderating role of task complexity. In R. M. Manchón (Ed.), *Writing and language learning: Advancing research agendas* (pp. 183–206). John Benjamins. <https://doi.org/10.1075/llt.56.08vas>
- Youn, S. J. (2013). *Validating task-based assessment of L2 pragmatics in interaction using mixed methods* [PhD Dissertation, University of Hawai'i at Manoa]. ProQuest Dissertations Publishing. 3577270. <https://www.proquest.com/openview/c52d8908a1efb15fd2b4fe203c2bf869/1?pq-origsite=gscholar&cbl=18750&diss=y>

Gavin BUI, PhD, is Associate Professor of Applied Linguistics at the Hang Seng University of Hong Kong. His research interests include task-based language teaching and third language learning motivation. His recent publications appeared in Language Teaching Research (2018), International Review of Applied Linguistics in Language Teaching (2019), System (2019), and Applied Linguistics Review (2020). He is co-editor of Asian Journal of English Language Teaching and president of the Hong Kong Association for Applied Linguistics.

Chi Him WONG is Senior Research Assistant in the Department of English at the Hang Seng University of Hong Kong. His research interests include English linguistics and Chinese phonology. He received his master's degree in Linguistics at the University College London.