Interactive Online and In-Person TESOL Part 2: Reporting Speech

Interactive Online and In-Person TESOL Part 2: Reporting Speech

David John WOOD

Abstract

Part 2 in the series **Interactive Online and In-person TESOL** *reviews the second-language (L2) acquisition CAF framework, complexity, accuracy and fluency. We take the system to task over its questionable use of terminology and commercial and literacy-based quantifying of acquisition. The distinct function of reported speech (RS) is located in a CAF framework, and its general discursive function is discussed. We present a methodology aimed at spontaneous interaction, with responses elicited by students among each other. This develops understanding and communication based on what students want to know and say, then reinforced by interactional reporting in a form of English comprehensible to speakers anywhere and at any level. Students increasingly used RES (Reported English Speech) to process, consolidate and develop their spoken communicative ability. The approach attempted a more practical version of CAF fluency, connecting and combining more effectively aspects of online and in-person EFL (English as a Foreign Language).*

1. Introduction

The CAF theoretical framework relates to previous studies (Wood, 2019, 2020a and 2020b). In addition, there are many studies of CAF features, especially in the context of task-based learning, and the way that certain task design features influence the development of each of the three components. RS, or Reported Speech, is often considered to be part of complexity in its form, but the application we discuss has broader implications. Exploring the relationship between particular communicative tasks (the inputs provided and the ways that they can affect the development of one or more of the CAF features) is another focus. Before considering such research's merits, we

scrutinize how high-frequency terminology is customized in EFL research, as it risks clouding and compromising communication.

Many studies just assume that their hybrid language use is self-explanatory. To redress the balance, we consider the limits of their meanings. We examine if RS is pedagogically worthwhile in terms of its effects on spoken communication ability in general. We consider its place in the CAF construct, and look at indicators relating to the effects of RS study, such as determining how much, if at all, students' RS ability improved, and their reactions to the methodology from various viewpoints.

2. Complexity, Accuracy and Fluency (CAF)

Basic perspectives of the CAF model are included to assess aspects of its validity, but years of CAF research make it impossible to discuss all applications at once. Semantic frequency suggests that fluency is the most important component, while accuracy and fluency are dependent at best. In research, CAF may often mean:

Complexity: described as the size, elaborateness, richness, and diversity of the learner's linguistic second language system (e.g., Housen & Kuiken, 2009);

Accuracy: defined as the degree of deviancy from a particular norm; deviations being usually characterized as errors (Wolfe-Quintero et al., 1998); and,

Fluency: explained as the ease, eloquence, and smoothness of speech or writing (Chambers, 1997, amongst many others, in Michel, 2017).

Each definition is tricky. The "norms" of accuracy are irreconcilably fluid, while equating distinct fluency skills and making "eloquence" crucial are both irrational, illustrating some odd notions the terminology can involve. Despite its inference in **Complexity**, in comparison to second-language fluency, there is little interest in or study of any learner's L1 (first language) fluency. Most searches of "L1" nearly always lead to English, thus demoting the relevance of learners' languages.

This seems a relic of linguistic imperialism. English L1 speakers are different to those of other languages; an obvious fact, but one often ignored. Past wealth and power will still dictate one attitude obstructing FLA (foreign-language-acquisition) understanding, if we assume everyone else must acquire English. The underlying assumption seems to be: L1 English speakers do not need to know much about learners' own linguistic frameworks, and even less about their individualized cultural identities. These examples suggest communication is not always a priority.

While CUP's dictionary (2022) definition of fluency sounds impartial ("the ability to speak or write a language easily, well and quickly") OUP's (2022, also online) immediately falls into the elit-

ist mindset of dictating that fluency is "the *quality* of being able to speak or *write* a language, especially a foreign language, easily and well" (my italics). As a significant percentage of the world is still learning basic oral English, assuming that literacy-based eloquence has greater relevance to L2s than L1s seems a definition based on an archaic English L1 mindset, or a bilingual one that includes English. Nor is CUP's definition particularly accurate. Like OUP, it equates spoken with written fluency, but the former must take precedent, while the latter is less essential. This literacy bias is unjustified. With almost anyone's L1 acquisition, listening and speaking precede, and are thus more essential than, reading and writing (just as reading is a prerequisite to writing). This reality seems overlooked. While EFL learners likely have literacy in their L1, it doesn't mean we can equalize or prioritize it in English as an L2. Many have managed to communicate without literacy for millennia. This is especially true when acquiring a new language.

The bias reinvents how the majority understand the term. The complexity, accuracy, fluency set (also sometimes reversed as FAC) has major face validity in the FLA international research community as an instrument to measure language proficiency. Although it indicates a positive attempt to systematize how we might better approach FLA teaching, its characterizations risk oversimplification. It is further compromised by high stakes commercial literacy testing (IELTS, TOEFL, and so on) not spoken communication, the essential first step for almost any L2.

We consider these descriptions (this writer's) to be more relevant and accurate:

Complexity is less essential to fluency than productive spoken communication. The latter allows (if not depends on) the readiness of communicators to explain complex utterances in simpler, reformulated synonyms. In fact, the most efficient communicators most often steer clear of complexity, or at least moderate and attune their levels to the lowest common denominator of respective participants.

Accuracy is dependent on and part of fluency. Communication will likely break down without basic accuracy, but it is still subordinate to fluency, and is neither an independent nor equal facet of communication, as native-speaker mistakes show.

Fluency in its most widely understood sense of communication, either embodies the other two facets, accuracy and fluency, or largely overrides them, as the prime element in initiating and sustaining basic interaction, especially listening fluency.

Similar to the research definitions above, the four-skill reduction of language - listening, speaking, reading and writing - also needs pinning down more rigorously across the board. The roles of other skills like communication risk being sidelined or ignored. Literacy (often redirected as literary studies in Japan) is vital in one's first language, but less important in any subsequent acquisition of other languages. *Mother tongue* suggests the language learners need most is oral. The Japanese for L1 *bokokugo* ("mother's country's language") implies a quite different concept. With

evolving computerized text translation, second language literacy's role seems less necessary than ever, but oracy cannot be subordinated. Until a learner has enough spoken ability to communicate, more complex study may just hold one back.

The order of priority for learners of EFL should still be oracy over literacy. Japan has long obsessed on literacy. Worse, the written language offered is often literary, archaic and accuracy-obsessed. This has compounded the wasted billions of yen spent in the vain attempt to achieve minimal fluency. Communication eludes millions in compulsory education now, as it has for decades. College English department course content remains passive written Japanese, transmitted by 90%+ teacher talk. Globally, spoken proficiency is more fundamental, even for native speakers.

Especially in the Japanese context of English education, teacher monologue takes up most of the scarce class time. Non-communicative, text-bound 'Conversation' accounts for 3% of College English majors' study or less in most curricula. Literature's perceived elite status tempts Japanese teachers merely to translate texts or sentences and explain them ad nauseum in Japanese, while grammar teachers focus on explaining moot rules, again, largely in Japanese. Literature and grammar are predominantly taught passively, defying any possibility of teaching English communication, and denying students any hope of legitimate acquisition. To emulate western institutions, colleges make literature and grammar their pillars.

The four skills (above) are regarded as equal in Japan and elsewhere (e.g., Columbia Gorge, 2022) meaning other equally if not more vital skills, not all linguistic or written, may go ignored. Limiting language puts too much emphasis on literacy, grammar and Japanese translation as priorities, obstructing skills more essential for communication. While the CAF model possibly "captures relevant aspects of L2 performance, a call for the inclusion of communicative or functional adequacy has been issued more than once in recent years" (Pallotti, 2009, in Michel, 2017). The latter defines the inclusion of communication into the equation as "the degree to which a learners' performance is more or less successful in achieving the task's goals efficiently scoring high on all 3 CAF measures can be communicatively inadequate and vice versa, which shows the independence of the two constructs. In language pedagogy and testing, communicative adequacy is one of the main goals" (Michel, ibid). CAF's communicative paradox is unresolved in research and teaching. It may still offer a tentative research framework for describing and identifying particular patterns in language performance as well as any of its few alternatives. However, the construct it suggests still requires a common-sensical balancing act within a perspective of a more widely acceptable meaning of communication and its achievement, which are the main aims of the methodology detailed in this study.

3. Task-based Complexity (TBC)

As with much EFL terminology (taught in non-English environments, in contrast to ESOL in areas that are predominantly English-speaking) semantic deflection is commonplace, whether intentional or otherwise, risking making English a foreign language even for those whose first language it is, and a kind of FL-within-an-FL for Japanese struggling with basic English communication. The meaning of "fluency" was challenged in the previous section. Here we confront terms like "tasks" and "task-based". In TEFL and TESOL research, and increasingly in practice, calling something a "task" seems more a smoke screen than a solution to major teaching problems. It is impossible to say exactly what a task is, or that it is a sole solution.

Outside the narrow world of TEFL research, a task might conjure up something mundane or arduous; a household chore or some difficult physical challenge. The 99% of native English speakers who are not teachers might find it hard to imagine it meant anything more than tidying or rote homework in a teaching context. This is even more so in Japan, where a western face or month's crash-course certificate have often been the only teaching qualifications required to be an EFL "teacher".

While task's original meaning (a tax or charge) may be lost to the past, one of the most widespread modern usages today is that of a computer mechanically carrying out predetermined, non-spontaneous actions. Between the two extremes lie various connotations, all foreign in meaning compared to high-frequency English. The latter is often confined to one's duty, or to an objective of almost any kind. This might be stretched to fit some notions of what goes on with the minority pursuing TEFL for research, but it is quite foreign to many teachers, and even more so to students.

Even if the term's application were completely satisfactory, at best it could mean anything done in class, rendering it a random term. A task might range from opening a textbook to mastering a language other than one's own. It is impossible to define where tasks begin or end, no matter how much research may attempt to remold the word, and many others like it. Anyone connected to language education, whether teachers or researchers, should use language transparently so that others can understand them in real-life terms. If not, then communication may be prevented before it can even begin. The language we use requires a common-sensical balancing within a perspective of more widely accepted understandings of communication, and the "tasks" we pursue or enforce in class should deliver results that are understandable to and instructive for students, and transparent to all.

Still, the application of the CAF model regarding the effects of task complexity needs to be considered, if for no other reason than to become more aware of what really helps in actual teach-

ing and all the irreconcilable EFL teaching situations that actually exist. According to Tsupa's research (2021) task-based complexity (TBC) may not be so significant: "Kuiken and Vedder (2012) looked into the impact of TBC on linguistic performance in oral and written modes and the impact of L2 proficiency on linguistic performance in different tasks. The researchers reported on 3 studies arguing the premises underlying the Cognition Hypothesis (Robinson, 2003) and the claims derived from it are problematic." This contrasts sharply with Anwad and Tavloki (2019) who stated: "Task complexity has a substantial influence on task performance." TBC has been defined as "attentional, memory, reasoning, and other information processing demands imposed by the structure of the task on the language learner" (Robinson, 2001, p. 29), or more simply put: "the cognitive load of a second language (L2) communication task" (Sasayama, 2016: 231).

Task complexity is assumed to affect L2 processing, production and acquisition, and used to "facilitate L2 production and acquisition, and how it interacts with psycholinguistic processes of attention allocation, noticing, and automaticity" (Tavloki 2018). The problem is less about applying terms to what we do than understanding why we choose to adopt one approach over another. This must be the first agenda item of any class for every teacher to discuss with students, who must be enabled to see why, and to respond fully in their L1, rarely the case in Japan.

Tavloki also asserts that intentional reasoning is significant. However, perhaps even more than in many other fields, in TEFL, linguistic and cultural factors raise major doubts about the possibility of objective measurement. Japanese students, for example, may exercise "intentional reasoning" less in their L1 than in many other cultures, as it is rarely demanded of them. Culturally, consensus is more valued.

It could also be that Japanese-style reasoning is not understood by TEFL researchers. Implying one culture's version of reasoning is better again implies cultural imperialism. Gender bias also exerts a greater influence than in cultures where English is the L1, as more assertiveness may exist there. The efficacy of so much research in the area of TBC is further brought into question as it favors non-spontaneous written modes rather than live oral communication. Even when the latter is attempted, the reliability of the methods used has been brought into question (e.g., Bachman, 1990 and Fulcher, 2014.) Bui's (2018) example claims:

"One wants learners to be able to perform with complex language, accurately, and fluently. (If) learners are challenged in any particular performance area, for example accuracy, task research can be relevant to pedagogical decisions. (In) acquisition, a developmental sequence (is) implied in the CAF analysis. New and more complex language is at first used haltingly. Next, greater control is achieved, with increasing accuracy. Then, as automatization takes place, greater fluency can be achieved, culminating in the new language being produced quickly, without error, and relatively effortlessly. CAF analysis is relevant to a sequence of acquisition and development."

(This author's under-linings and parentheses).

This perspective again seems questionable, due to the over-simplification of and assumptions behind such CAF research. Describing what should be about imperfect L2 communication in terms of perfect 'performance' conjures up connotations other than learners might expect in FL learning. Performance is less communicative than interaction. Ultimately, the TB approaches used appear more about testing and rote performance than real-life communication. The long, underlined sentence above might refer to a machine or some Garden of EFL Eden, being further removed than ever from down-to-Earth communication, which is human, sinful and imperfect.

The general sense of reliance on measurement in much CAF research is naturally central to teaching, but inevitably over-reliant on commercial, standardized written tests. Testing not only creates arbitrary results in many cases (despite what the merchants selling them assure people otherwise) it easily devolves into testing for the sake of testing, a vicious cycle of quasi-profiteering. Even more contentious is testing's overarching reliance on written forms over spoken ones, with one correct answer only being permissible, in direct contradiction to communication. By definition, communication is spoken and has no single or predeterminable answers.

4. Reported Speech RS

One focus of CAF studies is task complexity's role. RS study is an activity bridging complexity (an example of 'G3' advanced recursion at the highest-level-language grammars) and accuracy. It can increase class continuity and momentum if used as a homework task. The nature, aims and potential of its application merit more study, but RS does not receive enough attention in terms of its communicative potential, either in EFL research, or in actual classroom practice. The previous study (Wood, 2022) focused on the development of a single student's ability to communicate class interactions with reference to the developing use of reported speech.

In this second study, we look more closely at the English RES acquisition of 50 students in two separate classes. Both were conducted in mirror circumstances, under changing conditions, due to the recurring COVID-19 crises. Classes slated for in-person or online study were suddenly reversed, then changed back again. In the most extreme case, the switch occurred with only a single day's notice.

This disrupted basic communication, teaching and learning styles, as well as dismantling overall continuity in course planning and activation. Tech shortcomings and burgeoning administrative demands controlling education unwound an era lacking sufficient stability to guarantee the lesson-to-lesson requirements for effective acquisition, becoming insurmountable in many cases, with apathy reigning. Countering this, RES was a central common denominator for both

modes, steadying the imminent turmoil that might otherwise have wiped out the already limited chances for progress in a system heavy on commands and short on communication.

The role of RES in EFL is treated here in the wider context of teaching by talking, and precludes textbook use. This writer has not used textbooks to teach English communication throughout most of his decades-long English teaching career, believing that they stymy face-to-face spontaneity. In the Japanese EFL context, courses, whatever curricular names ascribed them, obsess on textbooks to the detriment of students' legitimate communicative needs and aspirations.

Therefore, unlike most English teaching in Japan, which deifies texts and pre-determined language, the method employed focused only on talking for communication. That required more interactive cooperation than textbooks offer. One way of achieving spontaneity is through live, spoken interaction about things that students think are really relevant to them. The teaching style discussed below, which forms the basis of this research, activates students to practice expressing themselves to learn about each other, one of communication's primary purposes.

To reinforce and realize the full potential of the activity, reported speech's role in the process is detailed. Rather than specifically aiming to increase CAF items, RES was used for communication. Fundamental to the technique is extending learning boundaries by recalling and summarizing class interactions for homework, as explained in the first study (2022). The process creates a perpetually self-regenerating cycle of study, reinforcing newly acquired abilities beyond students' immediate circumstances. This involves a process of conveying interaction contents in reported form, a seemingly complex skill, but one essential to communication.

This linking step helped students reflect on and extend their live English class experience, as well as developing communicative competence, by learning how to remind one other, or inform others, what they had said and heard. The development of this ability was studied in examples from the hundreds of reports generated by each course, comparing the respective gains of online and in-person formats. Sample recorded examples and analyses of online and in-person interactions are referred to below in order to show how they developed throughout a course. Students selected conversation topics of their own choice to ensure enthusiasm.

5. Some Qualities of Reported English Speech

First, we must ask some basic questions about Reported English Speech (RES):

What is RES? It includes recursive communication; what someone said to someone else when the person in question isn't present (and also about those present) and even to oneself, including feelings, as part of the English-speaking thought process.

How complex is it? Accurate expression requires various grammatical and syntactical

changes to the original utterances - there are simpler and more involved levels of reporting, including both statements and questions. It is most complex for speakers whose L1 does not include matching linguistic or cultural concepts.

How important is it to English communication? It is standard communication for native and bilingual speakers of English. Rather than giving a transcriptional report, as in a novel where dialogues are often conveyed in quotation marks, RES is a basic component of English speakers' linguistic identity using its own specific format.

Why is it difficult for learners of English? Depending on one's first language, the ways of expressing what we say may differ fundamentally. RES reflects a cultural concept, including inferring individual responsibility for one's statements. It is more common in older than younger speakers, so it's important for most EFL learners.

Here is an example indicating a fundamental point about this final facet for Japanese learners (Kajiki, 1957). The differences indicate that each language may have various unique linguistic and cultural systems for dealing with the concept:

#In Actual English Speech (AES) Kajiki quotes: "He said, 'I am right.'"

#In Reported Speech (RES) in English, Kajiki only quotes: "He said that he was right."

Kajiki doesn't consider either reported questions or incorporating them into reported interactions, but reported questions too are also fundamental to English.

Let's take for example something as simple as:

#AES: "How are you?" #RES: "He asked me how I was."

In the AES example above, and in all others in Kajiki's text, both sentences are translated in Japanese RJS identically (*kare wa jibun ga tadashii to omotta* or, *he/self/right/so/thought*). RJS and AJS are not differentiated the same way. While this may seem pedantic to English speakers' ears, it is standard in Japanese. The second simpler example is even more problematic as it is usually reported:

Genki ka to kiita = well (question indicator)*/so/asked* (leaving subjects ambiguous) or *O-genki to kiita =* (honorific 'O' indicating the person RJS is reported to) *well/so/asked*. For Japanese, figuring out subjects may be easier, but for English speakers, it seems too vague. Without the implicit detachment, the reporter may be considered to agree with the statement. However, in English, the possibility of not being right can be suggested by RES, and further intimated by facial expression or vocal tone to indicate the reporter doubts what had been said. While teaching the concept may not seem of vital importance, trying to sidestep or ignore it may risk incompleteness and even breakdowns in communication. More recent analyses of the concept in Japanese point to a form of RJS, but the pattern is rare. Yasui, M., (38 th W. Coast Formal Linguistics 2021, 483-492) uses *no da* (meaning 'because') as a sole example. However, the usage seems more limited in

terms of the connotations it can imply. In addition, it is not as high frequency as English RES, and the verb changes are different as Japanese don't express tenses in the same way, especially in the past and the future tenses. Syntactical changes such as with reported questions are also different as the mood of *no da* excludes RES questions.

Students not used to RES lacked practice and even the concept. Many students only used dialogue form, as in a novel or script. RES is distinct from the written mode of discursive transmission as it is also used widely in many spoken contexts. Native speakers therefore tend to use this form with less emphasis on 100% accurate grammar. Globally they use the dialog reporting form also. It is familiar, casual, and creates dramatic effect. RES is more complex than dialog script style because of the changes in tense and subject, but it is still a more standard format as it is more efficient for summarizing interactions, and in that way, lends itself to speedier communication, in terms of speedier delivery. While dialogue form suffices in certain situations, RES is essential for regular English communication. Learners who can't use it may seem less adept than those who can, even unable to function. Learners must be able to manage RES receptively to function in English contexts. One factor in the poor image of Japanese English use is RES unfamiliarity.

6. Research Overview

Continuing COVID-19-affected teaching has only increased the importance of effective methodology offering some transferability from online to in-person. Various teaching-environment configurations made third-year university students in two spoken communication classes a valid example to learn from. The preceding year's experience enabled improvements in aims, scope and so on. Factors included: more familiarity with online requirements and avoiding some of the easier pitfalls. This allowed stabler and better-quality analyses. On the other hand, the fluctuations of emotion triggered by a continuing pendulum of hope and despair (caused by the repeatedly changing teaching and learning modes usually imposed at short notice) still threatened the flow of effective teaching and acquisition.

Abstract research for its own sake would not be attractive to most students, so achieving valid and valuable communication goals was the sole priority. The effect of the latter would only be apparent in students' reactions and evaluations of their classes, not as mere research, but as helpful, bona-fide educational experiences. Using one line of inquiry from the preceding study (Wood, 2022) this second investigation considered the acquisition of reported speech ability more rigorously in 50 regularly attending students divided into two separate classes over two courses of ten core 90-minute classes, in 2 semesters in 2021, from May to July, then October to December. The first class scored higher on TOEIC than the second.

7. Class Outcomes and Descriptions

A preliminary summary of competent reported speech use is the clearest starting point. As each class was their only spoken communication class in their semester, it was possible to isolate students' progress, and thus to better evaluate the results:

<u>Group 1</u> 90%+ online compulsory speaking classes TOEIC 500+/CEFR B1:

    25 students, 10 x 90-minute weekly classes May to June 2021

    <u>Classes 1-3</u>: Average result: Correct RES 15; Incorrect 10; 60% accuracy

    <u>Classes 12-15</u>: Average result: Correct RES 23; Incorrect RS 2; =92% accuracy

    36% (8 out of 25) acquired RES, rising by 32% overall

<u>Group 2</u> 90%+ in-person compulsory speaking classes TOEIC 400+/CEFR A2:

    25 students,10 x 90-minute weekly classes October to December 2021

    <u>Classes 1-3</u>: Average result: Correct RES 5; Incorrect RS 20; 20% accuracy

    <u>Classes 12-15</u>: Average result: Correct RES 13; Incorrect 12; =52% accuracy

    36% (8 out of 25) acquired RES, rising by 32% overall

Reported speech was not the central focus of either group's classes. It was explained as being an element of English communication in the first class, and an optional goal to avoid extra anxiety. The rationale, in line with the previous sections, was that it made communicating in English more natural and comprehensible. Initial and final classes were especially confusing periods for students as teaching modes swung back and forth from online to in-person unpredictably with little notice. Students were placed in groups based on the previous year's TOEIC scores.

Accordingly, as the first group scored more highly, they were more adept at features of English like RS. About 60% began the course with basic RES command. Only 20% of the second group had any initial command. While the numbers of students who could acquire it obviously differed, there was some parity in the acquisition numbers. Thus, in both classes, 32% acquired facility in RS by the end of their respective courses. Both classes were conducted similarly, but the former, after one in-person lesson, went online until the end of the semester. Though slated as in-person, the latter started online then changed to in-person, and finally back again to online. No amount of planning could cope with such confusion. The overall total of 50 students registered for the two classes (whose progress was indicated above) all completed 90% or more of the 10 weekly RS assignments over the 3 months' class work, half predominantly online and half in-person. Each week, 2 or 3 students took turns to answer a variety of questions all others were assigned to think of before each class. Questions could not be used again in subsequent classes, and were split into two main types, either beginning with a verb or a question to encourage a wide use of English. The aim was for students to share and learn from each other their individual English, interests and in-

formation about themselves. Personal questions were avoided. Students could decline questions deemed sensitive. After each class, students were required to report the 2 or 3 students' answers to ten of the questions. Some were complex, involving spontaneous follow-up questions. Students received examples of how to summarize information and keep content culturally transparent. They had to explain any references to culture, places or proper names so that English speakers anywhere could understand. To this end, emphasis was on the use of high-frequency, internationally comprehensible vocabulary. References to individual or national ideas and culture hard to understand outside the immediate context had to be explained as clearly as possible. Students were to summarize and reorder content, not take dictation.

8. RES Sample Analysis

In the more advanced group, those not displaying RES ability initially had nearly all progressed by mid-course, but confusion still arose in 2 of the 10 who couldn't use RES at first. For example, instead of using the third person singular/plural (he or she/they) to indicate the reference was to anyone being spoken about, first person pronouns were still being used. The second group was far less used to the concept. Initially, most students tended to write out the question and answer as a dialogue, like a straight dictation exercise, although this was discouraged. Those attempting RES often retained the first-person pronoun, and several didn't reorder the questions from their originally inverted forms. In November, there was some increase in standard usage in Group 2, but most still displayed limited RES proficiency. The number writing the interactions as dialogue decreased slightly. The number attempting RES increased, but with persistent flaws. Nearly all students dropped the dialogue-style reports. The number of first rather than third-person references also dropped. Back-shifting tenses to agree with the head verb ('said', 'asked' and so on followed by a verb in the past tense) was still a problem for some. The gradual improvement continued in December, with some exceptions. Some could not break out of the dialogue-writing mode. However, most students demonstrated significantly increased facility, though back-shifting present to past was too difficult for about a third. Here are the early, mid and late-stage examples of a student from group 2:

October: *What is your hobby? My hobbies are playing basketball and chess.*

November: *What are you planning to do on the weekend?*

Answer: *Two students are working part-time and the other is going shopping.*

December: *Someone asked them what their club activities were when they were in high school. Two students said they were in a brass band and another student said she was playing badminton.* (Example of grasping the form of RS.)

9. Relevant Indicators

The previous section indicated that acquisition occurred in both groups, the first, primarily online, and the second, primarily in-person. There was some parity in the degree of progress, but various conditions as well as extra indicators need considering to grasp the approach's impact, and which mode of learning, online or in-person, may have been more effective, if either. In both modes, just over a third of students initially displaying no aptitude for RES managed to gain this vital basic skill, important for functioning and communicating smoothly anywhere in everyday English. While 60% of the higher-level class already had basic RES ability to begin with, only 20% of the second group had any initially. Obviously, the scope for improvement was therefore greater with the second group. This was balanced out to some extent by the fact that the general language proficiency of the latter group was assessed on TOEIC Listening as considerably lower, so less progress was likelier. The result seems to indicate that, although the actual gain in RES proficiency was equal in both groups at 36% or 8 out of 20 students in each, Group 2 out-performed initial ratings. Another 3 indicators with relevance to the approach's effect were:

    # The two groups' immediately subsequent TOEIC Listening scores;

    # Formal school-wide class evaluations; and,

    # Students' long-term reactions to their classes.

    # The two groups' immediately subsequent TOEIC Listening scores

As for TOEIC (the most widely used international test of English in Japan) Group 1 showed modest gains in the listening sections of the test administered after their course, as compared to the previous semester-end results. In contrast, Group 2's results rose significantly in a comparative sampling. This further suggests that, while both modes had a positive effect on listening ability, in-person teaching may have been the more effective, as only Group 2 had sustained face-to-face classes.

    # Formal school-wide class evaluations

While Group 1 student evaluations at the end of online classes were positive (Wood, 2022) Group 2's reactions were even more so. Though this may be due to some extent to the relief of returning to in-person classes, the turmoil of recurrent last-minute changes did nothing to dampen enthusiasm amidst repeated waves of COVID-19 crises. The comments of the former were discussed in the previous study. Here we focus on reactions to the most recent class. 80% of students responded in Group 2, an even higher figure than the 74% who replied in Group 1. (School-wide classes often receive only a single percentage response and, not uncommonly, zero.) School-wide comments are usually rare or negative and most evaluation questions do not relate directly to EFL. Of the 8 spoken classes in the same year group, the average response for the other 6 groups

of 161 students was 15%, 14% citing communication gain overall. As teachers swapped groups in semester 2, comparisons are valid. 74% & 80% of Groups 1 and 2, responded, 59% & 63% citing communication gain (similar to every year). Formal comments made were positive:

I enjoyed communicating in this class more than Reading and Writing classes.

We learnt how to use spoken English correctly to communicate.

Everyone learnt to speak accurately and communicate every time.

We acquired English communication skills by speaking and listening.

We learnt the ability to summarize what we said in English. (i.e., RES.)

This class was very useful to improve our English communication skills.

We learnt to enjoy communicating in English.

For the first time, I enjoyed communicating in English without any hesitation.

We acquired English naturally as we spoke and listened to each other every class.

I enjoyed everyone asking & answering our own questions each class.

Students' long-term reactions to their classes

The reaction of students taking the conversation courses to their final-year seminar application choice deserves noting. The 50 students researched in the two classes had 13 teachers' graduation seminars given by English Department teachers to choose from. Seminar titles included language, culture, literature and so on. These included 3 native speakers. 22 students from the groups, or **44%**, chose my spoken communication seminar, and especially those acquiring RES (see **11. Summary Table of Respective Gains**) while other teachers averaged only 2 to 3 applications or **5%**. In addition, only 2 students or **4%** chose the other native speakers, the remaining 26 choices being divided between the 10 department teachers who are Japanese. The choice was well-informed as, not only had these 50 students all taken my and sevral other native-speakers' classes in their 3 years of study, but all students received a 90-minute lecture from each department teacher on their seminars between October and December.

While there are many varieties of teaching English, being chosen by 44% of the 50 students surveyed with such a wide variety of 13 fully qualified and well experienced alternative teachers to choose from, indicates that this style of teaching was what students wanted the most. This year's seminar student's class evaluations (2021 to 2022) also vouched to satisfaction with the approach., which is spoken. Neither was this a one-time-only response as yearly applications often approach or exceed 20 students. There were significantly more respondents' comment responses than the school-wide average, which often sink to 1 or none:

I'm really happy I chose this seminar.

My communication matured and my motivation increased.

It was a marvelous class.

This seminar doesn't need any textbook.

The class focuses on talking and making someone understand you in English.

I thought this seminar was really important for me now, so I took part in this class.

The teacher taught us a lot of cultural differences between Japan and UK.

We learnt how to make everybody understand each other easily and logically.

We got everything that we need to have for the future and for globalization.

I think I've gotten international communication skills.

The happiest thing for me was that everyone was enthusiastic about learning.

I really enjoyed your seminar! Thank you so much!

10. Conclusions

After comparing hybrid CAF components' terminology in terms of their origins' most widely accepted meanings, we acknowledged the relevance to planning and enacting courses for EFL learners, with reservations. The issues left unaddressed, include the implication that native speakers are equally highly proficient. If they were all assessed in terms of the CAF research components, we could not explain language proficiency as anything except the property of professional interpreters.

What many teach and research under the name of EFL is skewed as classes and studies default to reading for grammar, not speaking for communication, influenced by written commercial language assessment and too many Japanese teachers' reluctance to speak or teach enough in the target language. The failure to use the language they base their status on to develop communication ability seems to be one reason why there is no significant development in students' English. The latter may be left with the idea they can or should avoid using English to communicate.

A standard 4-major-language skill set's questionable priorities, with listening and speaking regularly subordinated to reading and writing, were also brought into question. As the concept has been set in stone by educational authorities and Japanese practitioners without question, even the most enthusiastic students are left literally speechless. All the research in the world is useless if those responsible do not share spoken English more with their students. The notion of task complexity was also challenged as communicative relevance was seen as more important than trying to define levels of difficulty. Whatever actually works in whatever context seems more valid than blueprinting one system to cure every situation because of certain research, but the definition of success is unclear, and least researched.

We therefore sought to validate the reported-speech method employed. The distinct function of reported speech was discussed in CAF terms, and its general discursive and recursive nature

was mentioned in various contexts to indicate at least the need for receptive proficiency, as native English speakers use it so often. RES is perhaps most often considered as part of complexity, but the application discussed was more for its communicative potential. It aimed to reinforce class talk, as well as to connect the ongoing experience by extending content to each subsequent class in the form of short, rapid, globally transparent spoken recaps. RES also served the function of accuracy as an additional benefit.

The ultimate goal was to create confidence and outreach in communication. In addition to the obvious development of RES proficiency in the two classes studied (in different modes, online and in-person) external exam scores in listening indicated general improvement in students' receptive abilities. More qualitatively, official feedback from students on school-wide evaluations showed significantly more positive attitudes to the sole textless method, receiving over 4 times more responses, with most citing communication. Another indicator was the high application rate for students final and most important course, graduation seminars. The overall indication of RES development and acquisition from a combined-mode perspective strongly suggests that it is a technique positively applicable to both in-person and online EFL teaching and learning, with some likelihood that it may be more effectively teachable in-person. It also suited the random, hybrid switching which students had endured over the two years, like those studied in this paper.

This series will next aim to explore combining modes by researching what both students and teachers see as the merits of the two styles, separately and in combination, as their combined experience and reactions would likely reveal more than abstract and impersonal data. As COVID drifts on from year to year, more thorough, relevant and accessible research not yet in the public domain is needed.

11. Summary Table of Respective Gains

Criterion	Group 1 Online	Group 2 In person
RES-able students start>finish	15/25>23/25	5/25>13/25
RES acquisition	8/10	8/21
TOEIC Listening gains (sampled from students acquiring RES)	270>340=+70 TOEIC max=495	260>330=+70
Evaluation Response/Citing Communication *school averages*	**74%/59%** *13%/10%*	**18%/63%** *18%/12%*
Seminar application% (sampled from students acquiring RES)	100%	66%

The table suggests both groups made similar gains in RES, with improvements in listening ability as indicated by external examinations. Their respective evaluation responses were almost equally positive, at much higher rates than the 6 similar classes (**74% & 80%** to *13% & 18%*) as with those citing communication gains (**59% & 63%** to *10% & 12%*). As all teachers swapped classes in semester 2, the results comparison has considerable validity.

The difference in application rates among students acquiring RES capability indicates that fewer of Group 2 applied as they felt they had less chance of being accepted. Department rules dictate a cap of 11 students, meaning that the rejection rate was 50%, much higher than most other seminars. As this is a frequent tendency, and as students are aware of how difficult it is to qualify, the high application rate suggests an even higher approval rate.

References and Studies Cited

Anwad, A., & P. Tavloki (2018). Task complexity, language proficiency and working memory: Interaction effects on L2 speech performance. *International Review Applied Linguistics in LT*.

Bachman, L. (1990). *Fundamental Consideration in Language Testing*. OUP.

Bui, B., & P. Skehan (2018). In *TESOL ELT Encyclopedia*. Wiley & Sons, Inc.

Chambers (1997).

Council of Europe (1991). CEFR: Common European Framework of Reference.

Cambridge University Press (2022). Online English Dictionary.

Fulcher, G. (2014). *Testing second language speaking*. Routledge.

Housen, A., & F. Kuiken (2009). *Complexity, Accuracy, and Fluency in Second Language Acquisition*. OUP.

Housen, A., F. Kuiken, & Vedder, I. (2012). *Complexity, Accuracy, and Fluency: Definitions, Measurement and Research*. Benjamins.

Kajiki, R. *Eigonokiso* Obunsha, Tokyo (1957. Reprinted through 1989.)

Michel, M. (2017). *Complexity, Accuracy and Fluency in L1 Production*. (Routledge).

Oxford University Press (2022). Online English Dictionary.

Loewen, S., & M. Sato (2017). *Handbook of Instructed SLA*. Routledge.

Sasayama (2016). *Is a Task Really Complex?* Wiley Online.

Tsupa, Y. (2021). *Studies in Applied Linguistics & TESOL*. Teachers College, Columbia University, Vol. 21, No. 1, pp. 47-52 The Forum.

Wolfe-Quintero. K., S. Inagaki & H.Y. Kim (1998). *Second Language Development in Writing*. University of Hawai'i.

Wood, D.J. (2019, 2020 a & 2020 b). Low intensity MALL in a Japanese Context 1-3.

＿＿＿＿ (2022). Interactive Online and In-person TESOL Part 1.

(All Chikushijogakuen University in-house publications, Dazaifu City, Japan).

（デイビッド・ジョン・ウッド：英語学科　教授）

Interactive Online and In-Person TESOL Part 2: Reporting Speech

David John WOOD