JACET International Convention
Selected Papers
Vol. 7

ISSN 2188-8612

ISSN 2188-8612

JACET International Convention
Selected Papers
Volume 7

JACET 58th (2019) International Convention

Beyond ‘Borderless’: English Education in a Changing Society

August 28 – 30, 2019

Nagoya Institute of Technology

JACET International Convention Selected Papers, Volume 7
Published by the Japan Association of College English Teachers (JACET)

Cover Design
KASUNO, Shin-ichi

Editorial Office
The JACET Office
Address: 55 Yokotera-machi, Shinjuku-ku, Tokyo 162-0831 JAPAN
Phone: +81-3-3268-9686 / Fax: +81-3-3268-9695
E-mail: jacet@zb3.so-net.ne.jp

First published in September 2020

Contents

*** Plenary Lectures**

Daniel Perrin, Marlies Whitehouse. Beyond Disciplines, Domains, and Languages: Theory and Practice of Applied Linguistics in, with, and for a Changing Society (August 28)

Akira Tajino. ELT Research Revisited: A Soft Systems Approach (August 29)

Angel M. Y. Lin. Translanguaging and Multimodal Approaches to Content and Language Integrated Learning (CLIL): Innovating with the Multimodalities-Entextualization Cycle (MEC) (August 30)

Invited Papers

Beyond Disciplines, Domains, and Languages: Theory and Practice of Applied Linguistics in, With, and for a Changing Society

Daniel Perrin

Marlies Whitehouse

AILA

Abstract

This article explains how research "on" practitioners can be turned into research "for and with" practitioners (Cameron, Frazer, Rampton, & Richardson, 1992, p. 22) by including these practitioners in the research teams. Methodologically, it draws on two decades of multimethod research and knowledge transformation at the interface of Applied Linguistics and transdisciplinary action research on professional communication (Perrin, 2013). Empirically, it is based on large corpora of data collected in multilingual and multicultural workplaces. First, the article outlines transdisciplinary action research as a theoretical framework that enables researchers and practitioners to collaboratively develop sustainable solutions to real-world problems in which language use in general and text production in particular play a substantial role. Then, progression analysis is explained as a multimethod approach to investigate text production practices in natural environments such as workplaces. Examples from three domains (education, finance, and translation) illustrate what value transdisciplinary collaboration between academic researchers and practitioners can add to knowledge generation in Applied Linguistics. The article concludes by suggesting empirically based measures for research that contribute to the development of both theory and practice in Applied Linguistics.

Keywords: transdisciplinarity, education, translation, journalism, finance

Introduction: Applied Linguistics and Transdisciplinarity

In an increasingly globalized and interconnected world, communication entails the transgressing of boundaries between "discourse systems" (Scollon, Scollon, & Jones, 2012) such as linguistic varieties, natural languages, and entire semiotic systems used by discursive cultures rooted in regions, professions, and societal groups. Applied Linguists thus find themselves in growing demand not only inside but also outside the classroom. Society-at-large expects Applied Linguistics to identify and analyze socially relevant "practical problems of language and communication" (AILA, 2020) and to contribute to sustainable solutions. Such solutions are considered sustainable when they add long-term value from the perspectives of researchers, practitioners, and society-at-large (Padmanabhan, 2018).

In developing sustainable solutions, Applied Linguistics can draw on methodological knowledge generated in transdisciplinary research in general (Bernstein, 2015; Perrin & Kramsch, 2018, for a short historical overview) and in the research framework of transdisciplinary action research (TDA) in particular (e.g., Burns, 2016; Gustavsen, 2006; Pasmore, 2006; Reason & Bradbury, 2006; Stokols, 2006 for an introduction to action research in Applied Linguistics). TDA aims at facilitating theoretically grounded and systematic collaboration between researchers and practitioners, such as Applied Linguists on the one hand and teacher educators, financial analysts, translators, journalists, and policy makers, on the other.

From the perspective of academic researchers, preparing for such collaboration includes "opening up to and learning to listen to a much wider domain or people who can not only act as 'linguistic informants' and 'research subjects' but who can also become part of knowledge building" (Pennycook, 2018, p. 126). But listening to and communicating with people representing other

domains and their languages requires communicative and epistemological efforts. Not surprisingly, the methodological principles and practices of TDA have included, from the very beginning of TDA, language awareness as the key success factor of a systematic collaboration between practitioners and researchers in research teams (e.g., Klein, 2008, p. 407).

If this collaboration succeeds, the TDA research framework enables researchers and practitioners to jointly develop sustainable solutions to complex practical problems of, in our case, language use in general and text production in particular. In the next section, we describe the methods we[1] have used within the TDA framework to analyze and finally solve problems of written communication in increasingly multilingual and globally connected settings. Three examples, from the domains of education, finance, and translation, illustrate what it means to build knowledge in thorough collaboration with academic researchers and practitioners.

Methods: The Multi-Method Approach of Progression Analysis

In all of our research projects, we apply progression analysis, which is a multi-method approach that combines ethnographic observation and interviews, computer recording, and cue-based retrospective verbal protocols. Progression analysis has proven valuable in understanding the text production processes of practitioners such as journalists, communication professionals, financial analysts, and translators. It allows data to be obtained on three levels in order to investigate collaborative text production as a situated activity in organizational and societal frameworks.

Ethnographic Observation and Interviews

The first level of progression analysis considers the text producers (TPs) and the situation of text production, including the TPs' professional socialization and economic, institutional, and technological influences on the work situation as well as the specific task that the TPs must

accomplish. Data on the TPs' self-perceptions are obtained in semi-standardized interviews that focus on their professional experience and their workplaces. Ethnographic data are collected through unstructured participatory observations of organizational practices as well as interviews about them. Findings on this level include, for example, TPs' general language awareness in the area of coherence problems.

Computer Recording

The second level of progression analysis records every keystroke and writing movement in the emerging text with programs that run in the background behind the text editors that the TPs usually use, for instance behind the user interfaces of their company's editing systems. The recording can follow the writing process over several workstations and does not influence the performance of the editing system or the TP. The computer recordings provide information about what TPs actually do during the text production process, with every movement and revision step representing intermediate text versions in the production process. Findings on this level can reveal, for example, the activities that result in a specific text coherence problem.

Cue-Based Retrospective Verbal Protocol

The third level of progression analysis, the socio-cognitive conceptualization or reconstruction, draws on verbal data to infer the mental structures that might have guided the activities observed on the second level. After finishing a text production process, TPs view a playback of their process and watch as their text emerges. While doing so, they are prompted to continuously comment on what they did while writing. An audio recording is made of this verbalization and transcribed in a cue-based retrospective verbal protocol (RVP). The RVP is then encoded with respect to aspects of language awareness, writing strategies, and conscious practices. Findings on this level can provide

insights into, for example, a TP's conscious decisions that resulted in a coherence problem in his or her text.

In sum, progression analysis enables researchers to consider all of the revisions to the text as well as all of the electronic resources accessed during the production process; to trace the development of the emerging text; and finally, to reconstruct the TPs' considerations from different perspectives. The multiple levels of progression analysis allow the strategies and practices that TPs articulate in their cue-based retrospective verbalizations to be placed in relation to the situational analysis and the data from the computer recordings. Product features such as coherence problems in final texts become understandable as resulting from situated activity, i.e., as the results of complex text production activities in dynamic contexts.

Results: Examples From Three Domains of Writing

In this section, we outline how transdisciplinary research teams have used progression analysis to understand and improve text production in various domains and workplaces. In order to provide comparable examples across research projects and domains, we focus on one narrow subtopic of analysis: the coherence problems in evolving texts. By *coherence*, we understand the syntactic and semantic, as well as the explicit and implicit pragmatic connectivity within and across text elements of all sizes, ranging from single words to entire paragraphs, texts, and intertextual chains in discourse (e.g., Campbell, 1995; Crossley & McNamara, 2016).

Aspects of this phenomenon are illustrated[2] in the next three subsections with data from three projects: the MYMOMENT project that tracked children's essay writing in order to improve teacher education; the NATIONALBANK project, where the production of financial analysts' recommendations for investors was analyzed to improve stakeholder communication; and the CAPTURING TRANSLATION PROCESSES project with its focus on the use of information sources

and decision-making.

MYMOMENT: Tracking Writing Behavior to Improve Teacher Education

Children perform a variety of tasks using computers, with word, picture, and sound processing programs quickly becoming their natural text production environment. The development of keystroke logging programs has enabled researchers to track the process of text production without changing the environment for the TPs concerned. In a research project called MYMOMENT and its follow-up projects (e.g., Gnach, Wiesner, Bertschi-Kaufmann, & Perrin, 2007), hundreds of children in primary school grades one to five were provided with a web-based interactive environment for reading and writing stories and for making comments in class and at home. Text production processes were recorded automatically with progression analysis.

In the following example, the 4th-grader Doris (pseudonym) writes a German text entitled "Der Regenbogen" (English: The Rainbow) as a piece of free composition; she was able to determine both the form and the content herself. Extracts 1 to 3 show the production of the first 5 (of 30) sentences of Doris' text. The notation system used in the extract is called S-notation (Severinson-Eklundh & Kollberg, 1996), which marks insertions and deletions and indicates their sequence in the writing process. Wherever the writing is interrupted to delete or add something, S-notation inserts the break-character $|_n$ in the text. Deleted passages are in ^{n}[square brackets]^n and insertions in ^{n}{curly braces}^{n} , with the subscript and superscript numbers n indicating the order of these steps.

Doris writes the first two sentences fluently, immediately correcting typos (deletions 1 and 4) and making a conceptual change (deletion 2).

> Als es an einem Sonntag ^{1}[ndie$|_1$]^1 die ganze Zeit regnete und dann wieder die Sonne schien beschloss ich in den Wald zu gehen. ^{2}[Ich$|_2$]^2 Die frische Luft tat mir sehr g4[uet|4]4ut.

As it rained [1][nthe|1][1] the whole time one Sunday and then the sun shone again I decided to go for a walk in the wood. [2][I|2][2] The fresh air did me a lot of g[4][oed|4][4]ood.

Extract 1. First two sentences of a 4th-grader's composition

She begins the third sentence by saying that she saw a lot of animals. Then she deletes the beginning of the sentence and writes the converse (i.e., she didn't see as many animals as last week; deletions 5 and 6, insertion 7 in Extract 2).

Ich sah [5][viele Tie|5][5] nicht so viele Tiere wie let[6][s][2]|[6][7][z][7]te Woche

I [5][saw many ani|5][5] didn't see as many animals as la[6][z][6]|[6][7][s][7]t week

Extract 2. Third sentence of the 4th-grader's composition

Once she writes the third sentence, Doris moves back through the text, correcting the spelling of a word (i.e., letste to letzte). She then continues to write about her experience in the woods (Extract 3). Doris then immediately deletes the last part of the previous sentence and makes what will become a significant turning point in her story (deletion 8).

Aber plötzlich stand ich wie angewurzelt stehen. Ich sah einen [8][toten Igel vor m|8][8] halb toten Igel vor mir

But suddenly I stood rooted. I saw a [8][dead hedgehog in front of m|8][8]half dead hedgehog in front of me

Extract 3. Fourth and fifth sentence of the 4th-grader's composition

The rest of the story is written in the same linear way. The reader learns that the narrator wanted to take the hedgehog home but then decided to leave it because she was afraid of her mother's reaction. While going home under a rainbow, she felt sad and guilty. But when she came back the next day, she saw the hedgehog alive, fully recovered—which was, in her understanding, what the

rainbow had promised.

This story only works because Doris changed the description of the hedgehog from dead to half dead. As it turns out, this local change ensures the dramaturgical coherence of the narration. The girl's decisive conceptual change in revision 8 illustrates epistemic writing: typing is used as a means to understand what she wants (and does not want) to say in her text to make it a coherent story. But in contrast to many adult experienced TPs' behavior, global coherence seems to be established on the fly, while typing, not by planning key elements in advance. The 4th-grader Doris tells her story in a linear way, correcting typos and altering far-reaching dramaturgical decisions by deleting all the characters on her way back to the stretch she wants to change.

All in all, the MYMOMENT project suggests that analyses of text production processes can provide teacher education with detailed empirical information about how children at specific stages of development think, write, and, in this particular case, ensure narrative coherence—and how they are, by doing so, influenced by their writing environment, peers' feedback, and teachers' instructions. The approach fosters both an age-specific understanding of essay writing and the design of writing education in central questions such as how to establish narrative coherence in a text. However, such data recordings and analyses require thorough— transdisciplinary—collaboration between children, parents, teachers, and researchers throughout the research projects.

NATIONALBANK: Analyzing Financial Analysts' Recommendations for Investors

Another strand of transdisciplinary projects investigates text production in the domain of finance (e.g., Whitehouse & Perrin, 2015). Because they are constantly writing recommendations for investors, financial analysts can be considered professional TPs without a professional TP's background—their professional education mostly focuses on technical knowledge about banking and finance and neglects language awareness and writing skills. Counteracting forces affect their text

production: the complex and dynamic environment of financial markets with volatile conditions requires immediate action and reaction, flawless processes, and writing practices that allow analysts to write, both, as experts for experts, and as translators for financial laypersons.

The challenging writing environment is reflected in the text products, e.g., coherence breaks (see below), and in consequence, the opinions differ on what texts of financial analysts should, can and do achieve. While banks, for example, see analysts' recommendations as the basis for investment decisions, studies show that the recommendations are often difficult to understand for retail investors with low financial literacy (Lusardi, 2019). In the NATIONALBANK project and its follow-up projects, analysts' text products, writing processes, and workflows in financial institutions such as banks have been investigated in order to raise individual and organizational language awareness and an orientation towards addressees' communicative needs.

In this article, we briefly highlight one particular outcome of this transdisciplinary research: the insight that coherence breaks in text products tend to emerge between phase shifts in writing processes (e.g., Perrin & Wildi, 2012). Put simply, when TPs switch from linear progression to jumping back and forth in the emerging text, they risk losing control over their text and its coherence. The correlation between phase shifts and coherence problems is so strong that the latter can be predicted from the former in the case of short writing processes such as the ones discussed here. In the following two examples, financial analysts had to write one-page reports for the annual report of their bank, which took them around one hour each.

In Figure 1, the two graphs illustrate writing processes with and without such phase shifts. In both graphs, the x-axis shows the order of deletions and insertions over time, from the beginning to the end of the writing process. The y-axis, in contrast, shows in which order these deletions and insertions appear in the final text, from the top to the bottom of the text product. The graph visualizes the progress of a TP working through an emerging text—how she or he moves back and forth in the

text produced so far while writing. We have therefore termed this type of visualization a progression graph (Perrin, 2003).

In the left graph of Figure 1, a straight line from the upper left to the lower right indicates a completely linear writing process. This means that the writer never went back in the text to change something further up. Instead, the "point of inscription" (Matsuhashi, 1987) always stayed at the end of the text written so far. It does not mean that there were no corrections; the author actually made 342 changes while writing. However, these changes were all made immediately after writing, as a result of visual and cognitive feedback from looking at the material screen or its mental representation while externalizing the text by typing it.

In the right graph, in contrast, shifts between more linear and more fragmentary phases (Fürer, 2018) are visible. The second half of the text, from revision 783 to the very last revision, 1095, is written in a more or less linear way, with minor jumps back and forth for changes within the same sentence or paragraph. Before, however, the author constantly jumps back and forth across the entire text so far to make fundamental changes. This is where text coherence problems tend to occur in short writing processes, which can be explained by the fact that after many fundamental changes within a short time authors tend to lose track of what they wanted to write and what they actually wrote.

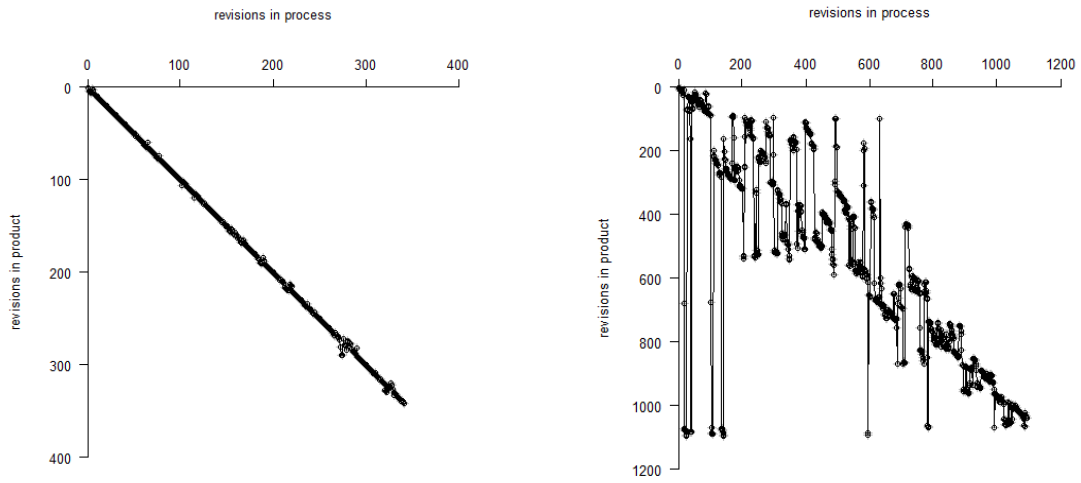


Figure 1. Linear progression (left) and multi-phase progression with phase shifts (right).

Such knowledge can help transdisciplinary research teams design measures for professional writing education for domain experts in institutions such as banks. The goal of the measures in the NATIONALBANK project was to enhance the comprehensibility and comprehensiveness of analysts' recommendations, among other ways by improving text coherence, in order to provide investors with a better basis for their decisions. Given the low average financial literacy of investors (e.g., Guiso & Viviano, 2013) and the (polito-)economic importance of investment decisions (Whitehouse, 2017), improving the communicative potential of analysts' recommendations is relevant not only in an economic but also in a societal context.

CAPTURING TRANSLATION PROCESSES: Revealing Translators' Use of Resources

In an extension to progression analysis, eye-tracking has been used in the laboratory to investigate at a micro level how TPs confirm or supplement their domain knowledge when working on texts in an unfamiliar area. The focus of attention can be tracked as TPs or translators access and read through digital sources and their emerging texts. This is what transdisciplinary projects such as

CAPTURING TRANSLATION PROCESSES do (e.g., Ehrensberger-Dow & Perrin, 2009; Ehrensberger-Dow & Perrin, 2013). In this project, we compared the translation practices of novices and experts in translation. The novices were selected from our university's professional education program, the experts from the staff of a research partner institution offering translation services to various businesses.

Figure 2 shows a screenshot from the eye-tracking software used in the lab. The connected spots indicate the foci and intensity of attention while a professional translator searched with Google for an acronym that appeared in a text about a military sonar detection system that had recently been banned in the US (see Massey & Ehrensberger-Dow, 2011). The translator was grappling with the task of adequately translating the acronym *MoD* for a German-language target readership that might be unaware of the differences between these two countries' military systems and its political relevance.

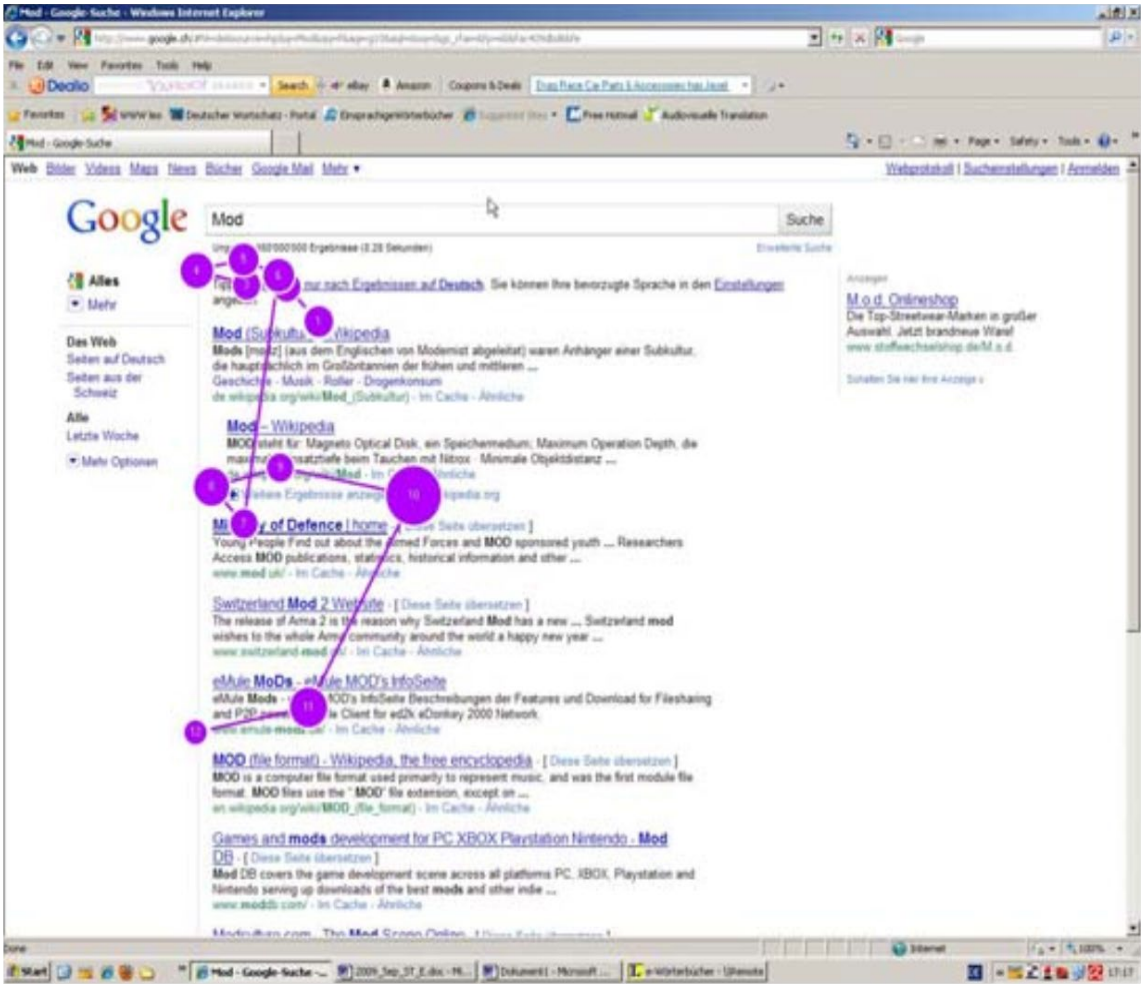


Figure 2. Eye-tracking visualization of a professional translator's use of digital resources.

The eye-tracking visualization shows that the translator skimmed through the first three hits, barely fixating on the descriptions in the snippet texts. The whole process took only 7 seconds before the translator returned to work on the emerging target text. In the subsequent RVP (see Methods, Cue-based retrospective verbal protocol), as he viewed his translation process, he remarked that he actually knew that *MoD* must mean *Ministry of Defence* in that context but that he had just wanted to check.

This purposeful confirmation by the professional is in stark contrast to the behavior of the large majority of student translators (no visualization in this paper). The ethnographic context analysis, on the first level of progression analysis (see Methods, Ethnographic observation and interviews) shows, not surprisingly, that students and novices have much less domain knowledge and experience than

professionals. In situations as the one shown above, they seem to have little idea of what to look for. In a case analyzed in detail elsewhere (Massey & Ehrensberger-Dow, 2011) even though the student translator's gaze falls on the solution to this particular translation problem in the list of hits that she checks, she fails to recognize it as such.

In this project, progression analysis has provided a framework that allows comparative investigations into the decision-making involved when translators shape their texts to meet the linguistic and cultural needs of their target audiences. Efficient exploitation of the appropriate digital resources is one of many indicators of translation expertise that have been identified in such research, which has fed into empirically based improvements in the education and professional development of translators (e.g., Massey & Ehrensberger-Dow, 2011). Again, such deep insights into professionals' workplace strategies and practices require these practitioners' thorough understanding and collaboration with the academics in transdisciplinary research projects and teams.

Discussion: Mediating the Relationship Between Disciplines and Real-Life Domains

In our research on, for, and with practitioners, we have analyzed "local" practices (Pennycook, 2010) of text production at workplaces in various domains, and we would like to present measures for systematic knowledge transformation between academic disciplines and real-life domains. They allow, for example, experienced practitioners' tacit knowledge to be made available to the entire organization, domain, and society at large. At the same time, these measures contribute to the development of theories of text production in, across, and beyond academic disciplines. The following catalogue groups the measures around the key conceptual aspects of TDA, of transdisciplinary action research: the *trans*-, the *action*, and the *research* aspect.

The *Trans*- Measures

Transdisciplinary action research transgresses and overcomes boundaries on four levels: between domains such as science, education, finance, and translation; between disciplines such as Applied Linguistics and communication studies; between institutions such as universities and banks; within individual researchers. This aspect of TDA is covered by the general methodological principle of integrating instead of excluding relevant stakeholders throughout a project. It is reflected in specific measures such as:

Getting support from the relevant parties. Instead of co-perpetuating the routines of institutions that structure "the ways in which research is funded, organized, conducted, and evaluated" (Pohl, Kerkhoff, Hirsch Hadorn, & Bammer, 2008, p. 417) and that are still strongly rooted in disciplinary paradigms (e.g., Choi & Pak, 2006; De Souza, 2017). In the MYMOMENT project, for example, all the parents of all the children in all the classes investigated needed to be made supporters of the project since they had to explicitly understand and agree to the recording of their children's text production activity. Of course, this applied to the pupils themselves as well as to the teachers, too. In addition, getting funding for such a project was a challenge at the time since it did not yet fit into funding institutions' schemes. Finally, the software company who had developed the online writing environment needed to be convinced that it was worth developing the keylogging modules and integrating them into the software.

Overcoming through emergent, creative approaches (e.g., Bernstein, 2015, pp. 16–17; Leavy, 2011). The apparent incompatibilites, for example of goals and timescales—instead of over-compromising, perhaps by sacrificing theoretical precision to a customer's interests in practical solutions (e.g., Hammersley, 2004) and "quick answers" (Agar, 2010, p. 8).

In the NATIONALBANK project, the partner from financial industries expected quick findings and absolute data privacy: The analyses of individual text progressions and collaborative workflows should immediately be transformed into as practical as possible guidelines and good practice cases

and both the data and the results should be kept under wraps forever. In contrast, the project partners from Applied Linguistics required time for comparative studies and theory building and aimed at discussing the study in their research communities. In the preparatory meetings of the project, the solution emerged to combine short and long timeframes. As a result, immediate findings were implemented on an ongoing basis as soon as possible after data gathering. Measures based on substantial generalizations, in contrast, were implemented after three years only. After ten years, the data were allowed to be shared with scientific communities.

Fostering communication and mutual learning. In a "reciprocal dialogue" (Jones & Stubbe, 2004, p.189) across stakeholder groups and project partners while doing Applied Linguistics as "the practice of language study" (Kramsch, 2015, p. 455)—instead of neglecting such communication and mutual learning, perhaps by failing to organize regular workshops for all the stakeholders during a research project (e.g., Morales, 2017, p. 33; Scholz & Steiner, 2015, Stokols, 2014, p. 56).

In the CAPTURING TRANSLATION PROCESSES project and its follow-ups, workshops for researchers, translation students, experienced professionals, and clients have been organized as platforms for knowledge transformations between all the relevant project phases. During these workshops, participants share their understandings of key concepts and make their mental reconstructions of the research objects explicit. It is mainly through finding words to share such reconstructions that the stakeholder groups become aware of differences in conceptualizations of the seemingly same world. Once accessible, such differences are excellent starting points for all the parties involved to broaden their horizons, think out of their boxes—and re-define in more inclusive ways what they are talking about, thinking of, and working with (e.g., Liddicoat, 2018).

In sum, trans- measures in TDA research ensure inclusion of all the relevant stakeholders throughout all the phases of a research project. This is in sharp contrast to disciplinary research, where researchers tend to develop study designs by keeping to themselves in what critics have named

the ivory tower or even "patterns of ignorance" (Rooney, 2011, p. 13), then access practitioners to gather data, then keep to themselves again to process this data, and only then contact the practical world in order to disseminate what they found out. TDA research actually starts with bringing the potential stakeholders together. All the research projects mentioned in the previous sections required such a substantial phase of collaborative preparation.

The *Action* Measures

TDA is oriented towards solving practical problems by taking action. Science tends to strive for "true" mid-range theories about situated activity; practitioners to look for "authentic" insights into their own circumstances and practices; and society at large to aim for "prudent" measures for solving practical problems (Kemmis, 1988, p. 46, based on Habermas, 1973). This aspect of TDA is covered by the general methodological principle of solving problems instead of shifting them. It is reflected in specific measures such as:

Observing and scrutinizing organizational power. For example beyond tendencies of "increasing professional effectiveness" (Argyris & Schön, 1974), together with all the stakeholder groups involved—instead of ignoring it, such as by unintentionally encouraging practitioners to amplify conflicts between management and employees (Jones & Stubbe, 2004).

Projects such as CAPTURING TRANSLATION PROCESSES involve stakeholders in the language and communication industries. Such companies tend to be organized hierarchically. This means that members in different roles at different levels of the organization have complementary tasks and pursue complementary or sometimes even contradictory goals. Whereas, in an oversimplified example, translators might want to keep to familiar routines, procedures, and standards, managers could intend to reshape the organization for more profit or growth. Both stakeholder groups then tend to interpret and reframe research in ways that foster what they have in

mind: The translators interpret the findings as supporting stability, the managers as supporting change. But there is more. Since TDA requires integrating practitioners from the very beginning of a project, the controversy needs to be made explicit and overcome before designing the research project and, in particular, before formulating the research question and goal. Integrating the previous knowledge of all the stakeholders is a substantial part of the "co-creation of knowledge" in TDA (Gravengaard, 2018).

Collaborating with practitioners and putting tacit knowledge to use (e.g., Padmanabhan, 2018; Stokols, 2014). Instead of exploiting them, perhaps as cheap sources of scientific data (Davies, 2007, p. 23) or destroying positive deviants' incentive to be one step ahead (T. McNamara, personal communication, July 5, 2010).

Discussing good practices as identified in projects such as MYMOMENT in classes of, for example, 4th-graders requires sensitivity to children's habits and norms of self-display in peer groups. On the one hand, children as well as adults may not want to share their successful practices with potential competitors. On the other hand, being identified as a good practitioner can be (mis-)understood as wanting to be better than the rest—which can result in ostracizing group dynamics. Moreover, including pupils, parents, and teachers in a TDA project such as MYMOMENT implies that these stakeholders learn something relevant about children's text production practices in general and about processes of establishing text coherence in particular—something above and beyond the practitioners' previous knowledge in this field and something that could not have been learnt as well in less complex, less demanding projects.

In sum, *action* measures in TDA research ensure sustainable solutions. In the projects outlined in the previous sections, the problems were solved by carefully mediating between the expectations of the stakeholders, by identifying crucial tacit knowledge at the bottom of the organization (where most of the situated activity of text production actually takes place), and by jointly implementing

knowledge transfer measures such as coaching and training.

The sustainable solutions based on projects such as NATIONALBANK include teaching, training, and coaching of financial analysts as text producers. Similar measures have been taken in the professional education of (business) journalists as the addressees of financial analyses. Furthermore, organizational development programs based on project findings have resulted in measurable improvement of the document cycling in many institutions in the finance industry. Finally, the results have impacted the public discourse about financial literacy and the academic reflection of this public discourse (Palmieri, Perrin, & Whitehouse, 2018).

The *Research* Measures

TDA, after all, is research, which is a theoretically based enterprise. In contrast to disciplinary research, TDA is oriented towards a "quadrangulation of disciplinary depth, multidisciplinary breadth, interdisciplinary integration, and transdisciplinary competencies" (Klein, 2008, p. 406). This aspect of TDA is covered by the general methodological principle of developing as theoretically and practically adequate as possible a multiperspective reconstruction of the interplay between layered structures and situated activity. It is reflected in specific measures such as:

Grasping the complexity of problems (e.g., Bernstein, 2015, p. 16). Instead of missing it, perhaps by limiting research to "the bleeding obvious" (Harcup, 2012) or by overlooking the diversity of scientific, professional, political, and societal views of questions, problems, and approaches (e.g., De Souza, 2017; Guo & Beckett, 2007; Kumaravadivelu, 2016).

In projects such as NATIONALBANK, identifying and questioning coherence problems in text products would fall short of theoretical and practical relevance. Both theorists and practitioners of financial communication can be expected to know that a lack of coherence risks lowering texts' communicative potential. In contrast, a theoretically interesting question could be why these

problems appear, whereas a practically relevant question could be who is most disadvantaged by these problems and how the problems can be avoided with context-sensitive measures. Similarly, applying readabilty formulae only to evaluate texts' communicative potential, as done in prominent studies in economics (e.g., Loughran & McDonald, 2014), fails to recognize that linguistics has developed more meaningful methods for evaluating text quality. In TDA frameworks, in dense collaboration across disciplines and domains, it is easier to become aware of the complexity of a problem and its potential solution —which, of course, makes it more demanding to actually develop this solution. However, solutions to practical problems can only be sustainable if the complexity of the problem has been taken into account.

Integrating practitioners' views (e.g., Roberts, 1997). In multiperspective approaches of bringing together "socially distributed" knowledge (Gibbons, 1994, p. 24) in "multi-stakeholder discourses" (Padmanabhan, 2018, p. 8) —instead of simply reproducing them, perhaps by uncritically reformulating everyday concepts in more complicated words (Bergman et al., 2010, p. 11; Kühl, 2008, p. 178).

The CAPTURING TRANSLATION PROCESSES project and its follow-up projects aimed at going beyond translators' professional knowledge and ethics. Translators as language professionals tend to be aware of a wide range of problems related to languages, language use, and language tools. As a result, educators and experts in the field of professional translation are often disappointed by translation studies. If research wants to add value from practitioners' perspectives, it has to understand what these practitioners already know—and to take this domain-specific shared knowledge into account when defining research questions and goals. This is why TDA postulates dense collaboration of scholars and field professionals from the very beginning of research projects (e.g., Makoni, 2003).

Exploiting methodological rigor. In favor of "mutual enhancement of disciplinary

epistemology" (Jantsch, 1970, p. 403; see also Padmanabhan, 2018, p. 10) —instead of losing it, perhaps by abandoning precise knowledge and rigorous methods (Arber, 1964; Denzin, 2010; Hirsch Hadorn et al., 2008; Jantsch, 1970, p. 403; Padmanabhan, 2018, pp. 9–10) of, e.g., linking case-specific information to abstract knowledge (Hammersley, 2004).

The MYMOMENT project exploited the large myMoment initiative in Swiss primary and secondary schools. This initiative resulted in a new writing ecosystem for 1st- to 9th-graders, including teacher training, essay writing classes, and an Internet-based environment to write essays and comment on the essays written by peers. What the research project added to this initiative is the recording part. Since all the essays of all the pupils were written by what later was termed *cloud computing*, it was comparably possible to track all the text production processes by implementing keylogging routines in the software on the host computer. As a result, such projects record the keyboard interaction of all the TPs for the duration of the project, which can easily result in tens of thousands of recorded text production processes after a year of recording. This data is minimally contextualized, for example with recording time and the TPs' log-in data. However, if research wants to understand why the TPs did what they did, it needs to overcome the big limitations of big-only data. This shift from big to deep data requires, for example, the theoretically sound selection of cases for cue-based retrospective verbal protocols (see Methods, Cue-based retrospective verbal protocol). In our projects, this is done following principles of grounded theory and theoretical sampling (e.g., Perrin, 2013, pp. 181–188).

In sum, research measures in TDA ensure the sound interplay of theory and practice. Ideally, this joint way of "coming to terms with reality" (Widdowson, 2001) results both in sustainable practical solutions *and* in advances of theory.

Conclusion: Mediating Between Realities

In all of our research projects, the findings from case studies were generalized according to principles of grounded theory to develop a model of the dynamic system of situated text production. This model contributes to both theory and practice in the field of writing by foregrounding the dynamics and complexity of collaborative text production. Most of the earlier models of writing and text production, in contrast, neglected aspects of collaboration since they had been developed in experimental settings where individual TPs were told to solve pre-defined problems in individual writing processes—which is quite the opposite of a text production task in natural professional settings (Perrin, 2013, pp. 150–152).

To put it simply, combining Applied Linguistics with principles and measures from transdisciplinary action research requires communication and collaboration across discursive cultures of stakeholders. Transdisciplinarity, in contrast to inter-disciplinarity, actually *is* about "mediating a relationship between two quite different planes of reality: that of the abstract discipline and that of the actual domains where the folk experience of language is to be found" (Widdowson, 2006, p.96). On the one hand, this raises project workloads and slows down research. On the other hand, TDA-informed research projects allow language-aware researchers to:

- enact their key competence of mediating between languages of academic and professional disciplines and their discursive cultures;
- provide evidence of their societal relevance by finding sustainable solutions to socially relevant problems in which language and communication play key roles;
- and contribute to the development of empirically grounded theories on language use in an increasingly complex, dynamic and interconnected world.

This means a threefold benefit, yielded by TDA in Applied Linguistics. Language matters, more than ever, and by taking appropriate measures, we are in an excellent position to do research that matters in real life—which is research on, for, and *with* practitioners.

Notes

1. The pronoun we refers to the research project groups as explained in this article. These research groups were and/or have been coordinated by the authors.

2. In all the three projects, data was collected and analyzed according to principles of grounded theory and theoretical sampling, allowing for generalizations to the level of mid-range theories (Perrin, 2013, pp. 181–188). Within their samples, the selected case studies represent relevant insights into the practices under research. Since the purpose of this article is not to discuss these mid-range theories but to explain what the TDA framework can add to AL-informed writing research, the generalization and generalizability of the findings from the selected case studies is not discussed in further details.

References

Agar, M. H. (2010). On the ethnographic part of the mix: A multi-genre tale of the field. *Organizational Research Methods, 13*(2), 286–303. doi:10.1177/1094428109340040

AILA. (2020). *What is AILA—mission and activities*. Retrieved from https://aila.info/about/

Arber, A. (1964). *The mind and the eye*. Cambridge: Cambridge University Press.

Argyris, C., & Schön, D. A. (1974). *Theory in practice: Increasing professional effectiveness*. San Francisco: Jossey-Bass.

Bergman, M. M., Eberle, T. S., Flick, U., Förster, T., Horber, E., Maeder, C., Mottier, V., Nadai, E., Rolshoven, J., Seale, C., & Widmer, J. (2010). *A statement on the meaning, quality assessment, and teaching of qualitative research methods*. Bern: Swiss Academy for Humanities and Social Sciences.

Bernstein, J. H. (2015). Transdisciplinarity: A review of its origins, development, and current issues. *Journal of Research Practice, 11*(1). Retrived from http://jrp.icaap.org/index.php/jrp/article/view/510/436

Burns, A. (2016). Action research. In J. D. Brown & C. Coombe (Eds.), *The Cambridge guide to research in language teaching and learning*. Cambridge: Cambridge University Press.

Cameron, D., Frazer, E., Rampton, B., & Richardson, K. (1992). *Researching language: Issues of power and method*. London: Routledge. doi:10.1017/S0272263100012729

Campbell, K. S. (1995). *Coherence, continuity and cohesion: Theoretical foundations for document design*. Hillsdale, NJ: Erlbaum.

Choi, B. C. K., & Pak, A. W. P. (2006). Multidisciplinarity, interdisciplinarity, and transdisciplinarity in health research, services, education, and policy: 1. Definitions, objectives, and evidence of effectiveness. *Clinical Investigative Medicine, 29*(6), 351–364.

Crossley, S. A., & McNamara, D. S. (2016). Say more and be more coherent: How text elaboration and cohesion can increase writing quality. *Journal of Writing Research, 7*(3), 351–370. doi:10.17239/jowr-2016.07.3.02

Davies, A. (2007). *An introduction to applied linguistics: From practice to theory* (2nd ed.). Edinburgh: Edinburgh University Press.

Denzin, N. K. (2010). Moments, mixed methods, and paradigm dialogs. *Qualitative Inquiry, 16*(419), 419–427. doi:10.1177/1077800410364608

De Souza, L. M. M. (2017). Epistemic diversity, lazy reason, and ethical translation in postcolonial contexts: The case of Indigenous educational policy in Brazil. In C. Kerfoot & K. Hyltenstam (Eds.), *Entangled discourses: South-north orders of visibility* (pp. 189–208). New York: Routledge.

Ehrensberger-Dow, M., & Perrin, D. (2009). Capturing translation processes to access metalinguistic awareness. *Across Languages and Cultures, 10*(2), 275–288. doi: 10.1556/Acr.10.2009.2.6

Ehrensberger-Dow, M., & Perrin, D. (2013). Applying a newswriting research approach to translation. *Target, 25*(1), 77–92. doi:10.1075/target.25.1.07ehr

Fürer, M. (2018). *Modeling, scaling and sequencing writing phases of Swiss television journalists* (Unpublished doctoral dissertation), University of Bern, Bern.

Gibbons, M. (1994). *The new production of knowledge: The dynamics of science and research in contemporary societies*. London: SAGE.

Gnach, A., Wiesner, E., Bertschi-Kaufmann, A., & Perrin, D. (2007). Children's writing processes when using computers: Insights based on combining analyses of product and process. *Research in Comparative and International Education, 2*(1), 13–28. Doi:10.23 04/rcie.2007.2.1.13

Gravengaard, G. (2018). Transforming knowledge: Cooperation with journalism's stakeholders. In C. Cotter & D. Perrin (Eds.), *Handbook of language and media* (pp. 489–504). London: Routledge.

Guiso, L., & Viviano, E. (2013). How much can financial literacy help? *Review of Finance, 19*(4), 1347–1382. doi:10.1093/rof/rfu033

Guo, Y., & Beckett, G. H. (2007). The hegemony of English as global language: Reclaiming local knowledge and culture in China. *Convergence, 40*(1–2), 117–132.

Gustavsen, B. (2006). Theory and practice: The mediating discourse. In P. Reason & H. Bradbury (Eds.), *Handbook of action research: Participative inquiry and practice* (pp. 17–26). London: Sage.

Habermas, J. (1973). *Theory and practice*. Boston: Beacon Press.

Hammersley, M. (2004). Action research: A contradiction in terms? *Oxford Review of Education, 30*(2), 165–181. doi:10.1080/0305498042000215502

Harcup, T. (2012). Questioning the 'bleeding obvious': What's the point of researching journalism? *Journalism, 13*(1), 21–37. doi:10.1177/1464884911400843

Hirsch Hadorn, G., Biber-Klemm, S., Grossenbacher-Mansuy, W., Hoffmann-Riem, H., Joyce, D., Pohl, C., Wiesmann, U., & Zemp, E. (2008). The emergence of transdisciplinarity as a form of research. In G. Hirsch Hadorn, H. Hoffmann-Riem, S. Biber-Klemm, W. Grossenbacher-Mansuy, D. Joye, C. Pohl, U. Wiesmann, & E. Zemp (Eds.), *Handbook of transdisciplinary research* (pp. 19–39). Berlin: Springer.

Jantsch, E. (1970). Inter- and transdisciplinary university: Systems approach to education and innovation. *Policy Sciences, 1*(4), 403–428. doi:10.1007/BF00145222

Jones, D., & Stubbe, M. (2004). Communication and the reflective practitioner: A shared perspective from sociolinguistics and organisational communication. *International Journal of Applied Linguistics, 14*(2), 185–211. doi:10.1111/j.1473-4192.2004.00059.x

Kemmis, S. (1988). Action research. In J. P. Keeves (Ed.), *Educational research methodology and measurement. An international handbook* (pp. 42–49). Oxford: Pergamon.

Klein, J. T. (2008). Education. In G. Hirsch Hadorn, H. Hoffmann-Riem, S. Biber-Klemm, W. Grossenbacher-Mansuy, D. Joye, C. Pohl, U. Wiesmann, & E. Zemp (Eds.), *Handbook of transdisciplinary research* (pp. 399–410). Berlin: Springer.

Kramsch, C. (2015). Applied linguistics: A theory of the practice. *Applied Linguistics, 36*(4), 454–465. doi:10.1093/applin/amv039

Kühl, S. (2008). *Coaching und Supervision. Zur personenorientierten Beratung in Organisationen*. Wiesbaden: Verlag für Sozialwissenschaften.

Kumaravadivelu, B. (2016). The decolonial option in English teaching: Can the subaltern act? *TESOL Quarterly, 50*(1), 66–85. doi:10.1002/tesq.202

Leavy, P. (2011). *Essentials of transdisciplinary research: Using problemcentered methodologies*. Walnut Creek: Left Coast.

Liddicoat, A. J. (2018). Language teaching and learning as a transdisciplinary endeavour: Multilingualism and epistemological diversity. *AILA Review, 2018*(31), 14–28. doi: 10.1075/aila.00011.lid

Loughran, T., & McDonald, B. D. (2014). Measuring readability in financial disclosures. *Journal of Finance, 69*(4), 1643–1671. doi:10.1111/jofi.12162

Lusardi, A. (2019). Financial literacy and the need for financial education: evidence and implications. *Swiss Journal of Economics and Statistics, 155*(1). doi:10.1186/s41937-019-0027-5

Makoni, S. (2003). Review: An introduction to applied linguistics: From practice to theory. *Applied Linguistics, 24*(1), 130–137. doi:10.1093/applin/24.1.130

Massey, G., & Ehrensberger-Dow, M. (2011). Technical and instrumental competence in the translator's workplace: Using process research to identify educational and ergonomic needs. *ILCEA Revue,* (14). doi:10.4000/ilcea.1060

Matsuhashi, A. (1987). Revising the text and altering the plan. In A. Matsuhashi (Ed.), *Writing in real time: Modeling text production processes* (pp. 197–223). Norwood: Ablex.

Morales, M. M. (2017). Creating the transdisciplinary individual: Guiding principles rooted in studio pedagogy. *Journal of Interdisciplinary Studies in Education, 6*(1), 28–42.

Padmanabhan, M. (2018). Introduction: Transdisciplinarity for sustainability. In M. Padmanabhan (Ed.), *Transdisciplinary research and sustainability. Collaboration, innovation and transformation* (pp. 1–32). London: Routledge.

Palmieri, R., Perrin, D., & Whitehouse, M. (2018). The pragmatics of financial communication. Part 1: From sources to the public sphere. *International Journal of Business Communication, 54*(4), 127–134. doi:10.1177/2329488418758449

Pasmore, W. (2006). Action reserach in the workplace: The socio-technical perspective. In P. Reason & H. Bradbury (Eds.), *Handbook of action research: Participative inquiry & practice* (pp. 38–48). London: Sage.

Pennycook, A. (2010). *Language as a local practice*. London: Routledge.

Pennycook, A. (2018). Applied linguistics as epistemic assemblage. *AILA Review, 31*(1), 113–134. doi:10.1075/aila.00015.pen

Perrin, D. (2003). Progression analysis (PA): Investigating writing strategies at the workplace. *Journal of Pragmatics, 35*(6), 907–921. doi:10.1016/S0378-2166(02)00125-X

Perrin, D. (2013). *The linguistics of newswriting*. Amsterdam et al.: John Benjamins. doi:10.1075/aals.11

Perrin, D., & Kramsch, C. (2018). Transdisciplinarity in applied linguistics. Introduction to the special issue. *AILA Review*, *31*, 1–13. doi:10.1075/aila.31

Perrin, D., & Wildi, M. (2012). Modeling writing phases. In M. Torrance, D. Alamargot, M. Castelló, F. Ganier, O. Kruse, A. Mangen, L. Tolchinsky, & L. van Waes (Eds.), *Learning to write effectively: Current trends in European research* (pp. 395–398). Bingley: Emerald. doi:10.1163/9781780529295_092

Pohl, C., Kerkhoff, L., Hirsch Hadorn, G., & Bammer, G. (2008). Integration. In G. Hirsch Hadorn, H. Hoffmann-Riem, S. Biber-Klemm, W. Grossenbacher-Mansuy, D. Joye, C. Pohl, U. Wiesmann, & E. Zemp (Eds.), *Handbook of transdisciplinary research* (pp. 411–424). Berlin: Springer.

Reason, P., & Bradbury, H. (2006). Introduction: Inquiry and participation in search of a world worthy of human aspiration. In P. Reason & H. Bradbury (Eds.), *Handbook of action research: Participative inquiry & practice* (pp. 2–14). London: Sage.

Roberts, C. (1997). There's nothing so practical as some good theories. *International Journal of Applied Linguistics, 7*(1), 66–78.

Rooney, P. (2011). The marginalization of feminist epistemology and what that reveals about epistemology 'proper'. In H. E. Grasswick (Ed.), *Feminist epistemology and philosophy of science: Power in knowledge* (pp. 3–24). Dordrecht: Springer. doi: 10.1007/978-1-4020-6835-5_1

Scholz, R. W., & Steiner, G. (2015). The real type and ideal type of transdisciplinary processes: Part II—What constraints and obstacles do we meet in practice? *Sustainability Science, 10*, 653–671. doi:10.1007/s11625-015-0327-3

Scollon, R., Scollon, S. W., & Jones, R. H. (2012). *Intercultural communication: A discourse approach* (3rd ed.). Cambridge: Wiley.

Severinson-Eklundh, K., & Kollberg, P. (1996). Computer tools for tracing the writing process: From keystroke records to S-notation. In G. Rijlaarsdam, H. Van den Bergh, & M. Couzijn (Eds.), *Current research in writing: Theories, models and methodology* (pp. 526–541). Amsterdam: Amsterdam University Press.

Stokols, D. (2006). Toward a science of transdisciplinary action research. *American Journal of Community Psychology, 38*(1), 63–77. doi:10.1007/s10464-006-9060-5

Stokols, D. (2014). Training the next generation of transdisciplinarians. In M. O'Rourke, S. Crowley, S. D. Eigenbrode, & J. D. Wulfhorst (Eds.), *Enhancing communication and collaboration in interdisciplinary research* (pp. 56–74). Los Angeles: Sage. doi: 10.4135/9781483352947.n4

Whitehouse, M. (2017). Financial analysts and their role in financial communication and investor relations. In A. V. Laskin (Ed.), *Handbook of financial communication and investor relations* (pp. 117–126). New York: Wiley. doi: 10.1002/9781119240822.ch10

Whitehouse, M., & Perrin, D. (2015). Comprehensibility and comprehensiveness of financial analysts' reports. *Studies in Communication Sciences, 15*(1), 111–119. doi:10.1016/j.scoms.2015.03.007

Widdowson, H. G. (2001). Coming to terms with reality: Applied linguistics in perspective. *AILA Review, 14*, 2–17.

Widdowson, H. G. (2006). Applied linguistics and interdisciplinarity. *International Journal of Applied Linguistics, 16*(1), 93–96. doi:10.1111/j.1473-4192.2006.00108.x

A Soft Systems Approach to ELT Research: Team Learning and Exploratory Practice

Akira Tajino

Nagoya University of Foreign Studies

Abstract

How can ELT research be further developed? This question is approached by examining the relationship between Team Learning, a new learning practice, and Exploratory Practice, an innovative form of practitioner research. The concept of Team Learning has emerged through a reinterpretation of a traditional Team Teaching practice by Soft Systems Methodology, a soft systems approach. The paper shows how Team Learning and Exploratory Practice support and develop each other through a shared set of common principles. The discussion about the commonalities of Team Learning and Exploratory Practice aims to illustrate how similar processes could be adapted to other approaches in ELT research and revitalize the field of Applied Linguistics with new dynamic practices.

Keywords: soft systems methodology, team teaching, team learning, exploratory practice

Introduction

Let us first ponder the following question: 'What does an international conference mean to us?' For some ELT professionals it means a place to present their research; for others it might be a place to exchange new ideas, opinions, and information with other participants; still others might think of

conferences as meeting places to make connections and collaborate with distant colleagues; or for some people it could be a place to learn about the latest resources and teaching materials.

Likewise, Team Teaching (hereafter, TT) can be viewed and interpreted in a number of ways. Traditionally, the *team* in TT has been limited to teachers. However, it is possible to include both teachers and students in the team when a wider view is taken. Learning as a team is an inclusive practice; i.e., students can take the role of the teacher, and the teacher that of the students. To reconceptualize the notion of *team*, Soft Systems Methodology (Checkland, 1983; 2001; Checkland & Scholes, 1999), a typical soft systems approach, was helpful as a tool to accommodate different worldviews and enhance the learning environment (Tajino & Tajino, 2000).

This paper will first discuss Team Learning (Tajino & Smith, 2016; Tajino & Tajino, 2000; Tajino & Walker, 1998) through the lens of SSM; second, see how it is related to Exploratory Practice (e.g., Allwright, 2003; Allwright & Hanks, 2009; Hanks, 2017); and finally, indicate that these practices can enhance collaboration and *quality of life* in the classroom. Therefore, when teachers join the students as equal partners in value-focused teams, it becomes possible to create a new dynamic wherein all participants become learners. This shift in role orientation allows all classroom participants to pursue aspects of the class that are important and meaningful to them. Out of this collaboration, it is possible for unforeseen emergent properties to arise and create an atmosphere for deeper understanding that could not otherwise be created in a traditional, teacher-centered classroom environment.

It is hoped that the processes described in this paper will inspire readers to explore new combinations of ELT concepts and expand the potential for new directions in ELT research.

Soft Systems Methodology

Soft Systems Methodology (SSM) is an action research approach developed by Peter

Checkland and his associates at Lancaster University in the 1970's in which complex situations are categorized and organized based on processes of enquiry and learning (Connell, 2001). SSM recognizes that individuals have independent perspectives, values and goals; and collaboration between individuals can be linked with logical goals resulting in an overall purpose creating complex systems and interactions. As these individually relevant perspectives are interwoven in human interactions, SSM focuses on the whole rather than individual parts, concluding that emergent properties of the context in question is indicative of the whole system. Emergent properties are:

> properties which are meaningful in relation to the whole entity, not in relation to its parts. The smell of ammonia is an emergent property of that gas which *has no meaning* at the level of the nitrogen and hydrogen which make up ammonia molecules.
>
> (Checkland, 1983, p. 669)

In thus being able to see emergent properties and the situation as a whole, people can understand the situation more holistically and take meaningful action to solve situations that are viewed as problematic within the context. While SSM is helpful to holistically understand various contexts from human relations to corporate organizations, of interest in this paper are problematic situations in ELT. Holliday (1990), for example, contends that SSM is valuable in order to understand where problems in curriculum development lie and also in structuring ethnographic research. If viewed as a tool for creating an environment of mutual understanding among all participants involved in a situation, the value of SSM becomes evident for understanding classroom life vis-à-vis the interactions between teachers and students.

To understand the context as a whole, SSM seeks to investigate each stakeholder's worldview and relationship with the context in question, and formulate practical solutions in problematic situations. SSM is claimed to be appropriate in "fuzzy ill-defined situations involving human beings and cultural considerations" (Checkland & Scholes, 1999, p. 10). Thus, SSM could be suitable for ELT research in which both students and teachers bring complex beliefs and goals into the classroom,

and negotiate knowledge sharing and inter-personal communication. Through the analysis of a complex, problematic situation in ELT, SSM can create manageable, actionable solutions to improve that situation.

According to Checkland and Scholes (1999), SSM is generally applied to situations in seven stages. These stages consist of two types of activities: real-world activities (Stages 1, 2, 5, 6, and 7) and systems thinking activities (Stages 3 and 4) listed below:

> **Stage 1.** Problem situation considered problematic
> **Stage 2.** Problem situation expressed
> **Stage 3.** Root definitions of relevant purposeful activity systems
> **Stage 4.** Conceptual models of the systems named in the root definitions
> **Stage 5.** Comparison of models and real world
> **Stage 6.** Changes: Systemically desirable and culturally feasible
> **Stage 7.** Action to improve the problem situation
>
> (Checkland & Scholes, 1999, p. 27)

In Stages 1 and 2, the model is used to identify problematic situations. In Stage 3, root definitions of relevant, purposeful human activities are formulated. Stage 4 builds conceptual models of these systems from the definitions in Stage 3, and Stage 5 compares them with the real world. Stage 6 establishes changes that are systemically desirable and culturally feasible, and finally Stage 7 implements actions to improve the situation.[1] It should be noted that, while all of these stages might be equally important for taking action, only the first four stages are relevant for the purpose of this study.

Critical to the root definitions and conceptual models in SSM is the transformational process. When considering the different factors in SSM, CATWOE (Customers, Actors, Transformation process, Weltanschauung (Worldview), Owners, and Environmental constraints) is a helpful set of analytical concepts:

> C: Customers: the victims or beneficiaries of transformation process (T)
> A: Actors: those who would do T
> T: Transformation process (T): the conversion of input to output
> W: Weltanschauung: the worldview which makes this T meaningful in context

O: Owner(s): those who could stop T
E: Environmental constraints: elements outside the system which it takes as given
(Checkland & Scholes, 1999, p. 35)

The concept of transformation process (T) is critical in this model. This situation can be understood in the classroom for ELT by envisioning the following context. For example, C represents a group of students; A is the primary teacher or multiple instructors; T is the interaction between student and teacher(s) regarding a topic of focus; W is the teaching approach used to structure the interaction in T; O includes other administrators at the school such as the principal; E is the workload of the teacher. The transformation process (T) is the target information or thought process that the teacher wants the students to acquire. It shapes the root definitions and conceptual models in SSM. The worldviews in this process (W) and *how* the information is perceived and transmitted greatly affect the education process. In this paper, no fundamental distinction is made between views and worldviews.

In recent years, SSM has been applied to and researched in multiple areas from private companies (Ferrari, Fares, & Martinelli, 2002) to health care services (Atkinson, Eldabi, & Paul, 2002; Checkland, 2001), but relevant to this context is SSM in English language pedagogy (Tajino, 2002; 2009; Tajino, James, & Kijima, 2005).[2] Due to SSM serving as a model to identify problematic situations and develop actionable solutions in complex circumstances, this model fits well into language pedagogy where there are multiple stakeholders (e.g., students, teachers, administrators) and multiple contexts (e.g., classrooms, curriculum, education goals). SSM can help to distill and unravel the complexity in ELT contexts, and develop actionable solutions from multiple potential vectors.

Team Teaching and Team Learning

The concept of Team Learning (TL) in ELT has emerged through a reflection on the team-

teaching practice in Japanese EFL contexts (see, for example, Tajino & Tajino, 2000). TT is a common pedagogic practice in ELT and other teaching contexts where two or more teachers come together to co-teach a class. This practice can be found at almost all levels of education from kindergarten to higher education, and is conducted in many countries. Particularly in Asian countries, co-taught English lessons are common, and the topic of TT has already been widely discussed in language education literature (Carless, 2006; Davidson, 2006; Gladman, 2015; Johannes, 2012; Liu, 2008; Luo, 2010; Ng, 2015; Park, 2014; Perry & Stewart, 2005; Sakamoto, 2020).

A review of TT in the Japanese context is a useful entry point to consider different versions of the concept. The most common place to see TT in Japan is in public elementary, junior, and senior high schools where a Japanese teacher of English (JTE) are often paired with an assistant language teacher (ALT)[3] who is usually a native or near-native English-speaking teacher. TT is generally defined as:

> a concerted endeavour made jointly by the JTE and the assistant English teacher AET in an English language classroom in which the students, the JTE and the AET are engaged in communicative activities.
>
> (Brumby & Wada, 1990, Introduction)

Through having both a JTE and an ALT in the classroom in Japan, it is possible that different teachers can offer multifaceted inputs to help students develop communicative competence and deepen their understanding of different cultures.

Research about TT in Japan has determined that the relationships and views of the teachers are key to determining synchronicity in the classroom. Different educational and cultural values, language barriers, and conflicting views of the classroom can lead to rifts between JTEs and ALTs (McConnell, 2000; Miyazato, 2009; Pearce, Stewart, & Tajino, 2019). Regarding worldviews, it was found that JTEs can perceive team-taught classrooms very differently. Some regard team taught lessons as a place where students feel tense or nervous, deepen their understanding of different

cultures, or have the opportunity to enjoy communicating in English (Tajino, 2002). Though the examples have been from TT research in Japan, it is not an exaggeration to say that teacher relationships or worldviews in any teaching context will impact how teachers interact and divide roles in the classroom.

Considering different relationships and worldviews among teachers in TT environments, Tajino and Tajino (2000) classified TT into two versions: a weak and a strong version. The weak version is an assembly of different roles in the classroom, where each teacher *cooperates* with the other yet plays a separate individual role. This can be described using a metaphor such as a musical concert where a pianist and a singer share the stage, but both play and sing at separate times.

The strong version posits that *collaboration* between the teachers will lead to a more dynamic class; in other words, both teachers work together to create an environment where the whole is not just the sum of its parts, thus setting the stage for emergent properties to arise (see Kato & Dalsky, 2019 for the difference in meaning between cooperation and collaboration). As Brumby and Wada (1990) state, TT should be a "concerted endeavour" (p. 3) made jointly by both teachers. In the strong version of TT, both teachers collaborate to achieve a "harmonious duet" (Tajino & Tajino, 2000, p. 6) in the classroom, not simply taking on separate roles, but collaborating in all aspects of the classroom experience. Continuing the above metaphor, the strong version of TT would be more akin to a duet where the sounds of both musicians merge and synchronize to create a sound both individually could not produce. A simplified descriptions using equations of both versions can be seen below:

Weak version of Team Teaching: X+Y
Teachers with separate, individual roles in the classroom

Strong version of Team Teaching: XY
Collaboration between teachers resonates to create a class where emergent properties arise.

At this point it would be worthwhile to consider the definition of *team* in the classroom and see

if it can be expanded further. One issue with traditional TT models is that teachers as a team only teach and students simply learn from the teacher team. In this model, students have little or no chance to speak and express themselves. Given that the objectives of ELT include the improvement of speech communication skills, this model would be insufficient.

Tajino and Tajino (2000) argue that TT can be reinterpreted as TL; i.e. teachers *as a team* can also listen to or learn from students (teachers-as-learners). This would provide the students with more opportunity to use and improve their productive skills such as speaking and writing skills (see Figure 1). This can be considered a narrow view of TL.

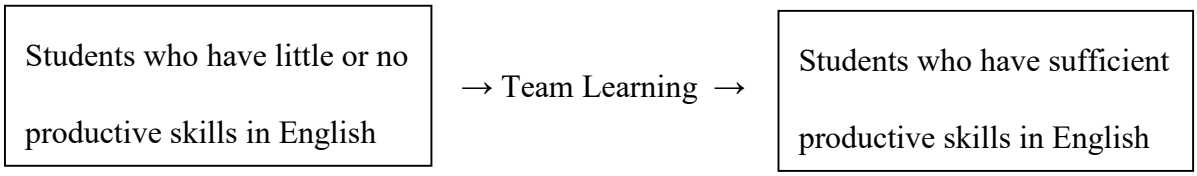


Figure 1. A transformation process model for a narrow view of Team Learning.

Allwright (2005a) claims that both "teachers and learners co-construct their lessons" (p. 16), thus it is important to include learners (as well as teachers) in the definition of *team*. Reconceptualizing the narrow view of TL to include the entire class as a team (both teachers and students together), opens up classroom instruction to incorporate a broader view of learning. This broadened discussion leads educators to new possibilities. With an expanded conceptualization of TL that combines the worldviews of the students and teachers to reimagine the concept of team, what is most likely for the further development of TL in ELT? The next section will describe the potential of TL to go beyond TT, and create more inclusive classrooms wherein all participants engage as a *harmonious orchestra*.

Team Learning: A Broad View

Unlike a traditional type of TT, TL values collaboration and sharing among all participants in the classroom, both teachers and students alike (Tajino & Tajino, 2000; Tajino & Walker, 1998). Traditional learning styles such as lecturing would often result in one-way transference of knowledge, either from the teacher to the learner, or in some cases interaction between students and teachers in the form of questions and answers. TL, on the other hand, involves all classroom participants in the learning process where the goal is sharing knowledge and understanding. In TL, the teacher–student dichotomy is discarded, and the teacher is viewed as a fellow team member with the students where the control of interaction takes place between and among students and teachers, allowing for deep, collaborative learning. Having all members in the class as equal partners allows for greater chances of collaboration, and the generation of emergent properties.

The concept of TL comes from Tajino and Walker (1998) and was developed in Tajino and Tajino's (2000) observation of team teachers in Japanese secondary schools. During this process, they saw the potential for expanding the concept of *team* beyond co-teachers. In Tajino and Smith (2016) it was argued that the concept of *team* should take a broad context in English teaching to include everybody participating in the lesson. Classes are structured to encourage *fluid roles* among all classroom participants, where content and knowledge are shared.

In order for teachers to leave their traditional stance in the classroom behind and become team participants, several key steps must be taken to promote *role sharing*, *role fluidity*, and *role synergy* (Stewart, Dalsky, & Tajino, 2019). To begin, teachers must accept their role as a learner and actively reduce the distance between themselves and the students, allowing for *role sharing*. Additionally, *role fluidity* must be promoted, where students must understand that their input is important and valued in the classroom. While training students to be more autonomous and take responsibility for their own learning, the teacher can also become a team member. This allows students to feel confident and trusted to take more central roles. Finally, *role synergy* occurs when all participants are

completely engaged in the learning process. This is called the "optimal state of engagement" and emerges when students "behave unselfconsciously as learners" (Stewart et al., 2019, p. 4). It is in this space between *role fluidity* and *role synergy* where TL occurs.

Through the practice of TL, and especially during *role synergy* where classroom participants are completely engaged as learners, mutual understanding and shared comprehension can produce *emergent properties*. As discussed earlier emergent properties result from a complex system where the whole is more than just the sum of its parts. Students can also learn more than what teachers teach.

> [...] what learners actually learn from a lesson is going to be less than all the teacher's teaching points, what they can and might learn from a lesson is also potentially, and perhaps normally, much richer than just the sum total of such teaching points.
>
> (Allwright, 2005a, p. 15)

Allwright and Bailey (1991) point out that both students and teachers bring their own sets of values and individual experiences to the classroom. Therefore, classroom life involves the interplay of these different values, beliefs and experiences, thus making the classroom a richer, more complex environment.

Through the merging of different ideas in this setting, a value-centered classroom environment can be created. In this context, *values* can be considered as beliefs about what is considered important in the language learning classroom. One such value could be the importance of communication with people in other cultures (Tajino & Smith, 2016). Another could be enhancement of quality of life in the classroom. The concept of *quality of life* will be discussed in detail in the following section, and the potential for synergy between Exploratory Practice and TL will become more apparent.

Exploratory Practice

Exploratory Practice (EP) is a practitioner-oriented research approach that focuses on bringing together all classroom participants to expand their understanding about particular *puzzles* they have to increase the quality of life in the classroom (see Hanks, 2019a). Hanks (2019b, p. 9) claims:

> developing understanding is prioritized as a guiding principle of Exploratory Practice, as Hanks (1999; 2009; 2017a) has elucidated, with the aim of encouraging curiosity-driven, practitioner-led research which inquires into contextually-appropriate puzzles set by the learners and teachers themselves.

EP includes all learners as integral to the learning process, and provides an opportunity to deepen their understanding of classroom life from either the student or teacher perspective. Allwirght (2005b) offers a succinct definition of EP:

> Exploratory Practice (EP) is an indefinitely sustainable way for classroom language teachers and learners, while getting on with their own learning and teaching, to develop their own understandings of life in the language classroom.
>
> (Allwright, 2005b, p. 361)

Greater understanding is reached by participants in lessons as they investigate aspects of their own learning (and teaching) that they find puzzling. A *puzzle* is not a problem, but rather a question that emerges during classroom interactions. The goal of exploring individual (or group) puzzles is to achieve some higher degree of understanding. It also often involves engagement in the process of reaching a consensus about the understandings of all participants. Thus, through the collaborative process of EP teachers and students alike can enhance their understandings of learning and elevate the quality of life in the classroom.

The key issues that guide EP focus on *what*, *who*, and *how*. *What* issues regard different aspects of classroom activities, and life in general among the learners. *Who* issues strive to involve all learners in working through the puzzle, and create a collaborative environment of mutual understanding. *How* issues represent logistical concerns about minimizing the workload, and creating continuous professional growth opportunities for learners involved. These principles are summarized

below. Various examples of the benefits of collaboration using EP have been found in recent studies (e.g., Hiratsuka, 2019; Rowland, 2011; Slimani-Rolls & Kiely, 2014; Smith, 2009).

Seven principles for inclusive practitioner research

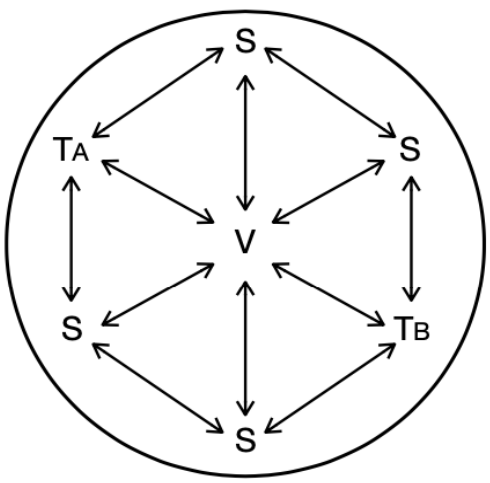
The 'what' issues:
1. Focus on *quality of life* as the fundamental issue.
2. Work to *understand* it, before thinking about solving problems.

The 'who' issues:
3. Involve *everybody* as practitioners developing their own understandings.
4. Work to bring people *together* in a common enterprise.
5. Work cooperatively for *mutual* development.

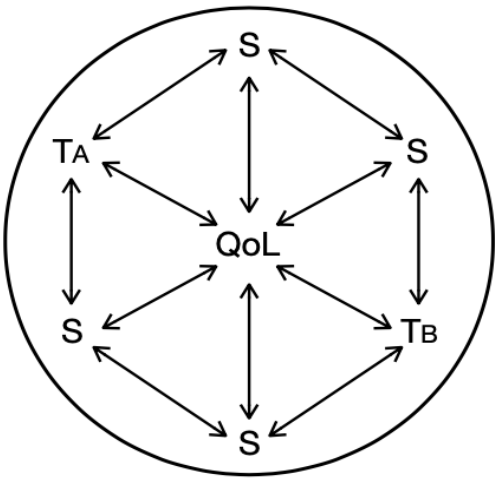
The 'how' issues:
6. Make it a *continuous* enterprise.
7. *Minimise the burden* by integrating the work for understanding into normal pedagogic practice.

(Allwright & Hanks, 2009, p. 260)

In this paper, the *what* issues are of particular interest as the goal of enhancing classroom quality of life parallels the value-centered nature of TL. Figures 2a & 2b illustrates how both *values* and *quality of life* are at the center of both practices (Tajino & Smith, 2016).



S: Student; T: Teacher



S: Student; T: Teacher

Figure 2a. [left]: A Value-centered Team Learning Model.
Figure 2b. [right]: An Exploratory Practice Model.

(Tajino & Smith, 2016, p. 20 & 22)

When viewed through the lens of SSM, TL and EP have great potential for collaboration and

emergent properties. In the next section, that potential will be explored.

Team Learning and Exploratory Practice

By providing a systemic framework, SSM can be used to promote TL and enhance EP. In this way, SSM opens possibilities to understand the *whys* that underlie the *whats* in classroom life through the transformation-process model (see Tajino, 2009).

Thus far in the paper, it has been demonstrated that both TL and EP have similar goals, but it was not made clear how they can combine in a mutually supportive way. As noted above, there are seven principles in EP that can be categorized into three issues: *what, who,* and *how*. This section will discuss the potential for collaboration between TL and EP in terms of these categories (see Tajino & Smith, 2016).

The 'what' issues

With quality of life as a focal point to foster understanding among classroom participants, instructors could use TL to conceptualize and facilitate this mutual cooperation to reach *understanding*. In this regard, activities used in TL such as discussion tasks to negotiate a consensus can be focused on how classroom life could be improved. Team building tasks create a necessary environment for EP to be effective, and facilitate the *understanding* necessary for meaningful classroom quality of life (see Stewart, 2018).

Beyond creating collaborative environments, student reflection about their own perceptions as well as sharing of classmate and teacher perceptions are also important. In order to facilitate reflection on the TL environment, students and teachers could hold small group evaluations with brief interviews at the mid- and end-points of lessons regarding their thoughts and feelings related to their learning experiences. Additionally, such interview information allows for the development of critical thinking skills and communicative competence. This practice could support quality of life by

having learners work to understand other team members' feelings while working through the puzzle.

The 'who' issues

The *who* issues focus on including all classroom learners (i.e., teachers and students) in the investigation of a puzzle, to foster mutual understanding and collaboration in the class.

One means of promoting such teamwork is to include both teachers and learners as practitioners in deciding how they will interact in the classroom. By having everyone set goals and decide evaluation practices, every learner can more easily participate in learning tasks. Evaluation practices could be created to assess both on an individual scale, and the team as a whole to create incentives to work together.

Beyond the norm of teachers simply deciding activity goals and creating evaluation rubrics, having the class collaboratively develop the tasks can further bring all participants together. Teachers who elect to implement this type of negotiated syllabus may find it useful to provide the class with a list of optional activities to choose from to promote team interaction.

The 'how' issues

TL may be beneficial for solving the logistical problems of using EP in the classroom. Considering that a very specific environment of mutual understanding and collaboration is necessary for EP to be effective, looking at practical examples of how TL has been integrated into existing curricula can offer guidance. TL aims to accept the constraints of the curriculum (e.g. preparation for entrance examinations) by working towards course goals in ways that allow students and teachers to also enhance the quality of life in their classroom.

Efforts should also be made to keep the class stable and meaningful. To this degree, practitioners in the classroom should continuously reflect on how well learning tasks are suited for the current objectives. This means that the validity of tasks should be considered both on the individual and team level, and be adapted accordingly to the needs of the situation.

Conclusion

This paper began with the question: 'What does an international conference mean to us?' In considering various reasons why people might join a conference, it is possible to re-conceptualize such an event. These new perspectives could add new meaning, leading to previously unconsidered directions that enhance the overall success of the conference. Soft Systems Methodology presupposes that different people have different views and interpretations of the same situation. Eliciting and making use of the range of views or worldviews opens new possibilities to improve the situation for everyone involved as they negotiate the meaning of it collaboratively.

This paper has discussed ways in which Soft Systems Methodology could be applied to ELT research. Through a soft systems approach to Team Teaching, two interpretations of the concept of *team* have been considered (i.e., X+Y and XY). In addition, two views of Team Learning have been elaborated on: a narrow and a broad view. The narrow view refers to the teacher team, when only teachers are considered in the team with students excluded, whereas the broad view considers all participants in the classroom as members of one *team*.

The value-centered model of Team Learning is not a concept in isolation. This paper has shown how Exploratory Practice shares common features with Team Learning through its focus on *quality of life*. Teachers who understand this shared focal point between Team Learning and Exploratory Practice can search for new ways to enhance classroom language learning. This is an exciting way for new knowledge to be created in the field of ELT. The discussion in this paper has been organized according to the interplay of Soft Systems Methodology, Team Learning, and Exploratory Practice to illustrate some possibilities for future directions in ELT research.

Acknowledgements

The author is grateful to Dick Allwright and Judith Hanks for their innovative and ground-breaking work on Exploratory Practice, and also to Yoshitaka Kato, John Andras Molnar, and Tim Stewart for their invaluable comments on earlier drafts of this paper.

Notes

1. This model is highly flexible, and it should be noted that not all the stages need to be used depending on the needs of the situation. For example, a four-stage model is also possible (see, for example, Checkland & Poulter, 2006).
2. SSM has increasingly received attention in foreign language pedagogy and has been applied to research areas and topics including classroom research (Kato & Dalsky, 2019; Stewart, 2019), ESP research (Maswana &Tajino, 2020; Terauchi & Maswana, 2019), grammar instruction (Tajino, 2019), instructional design (Hosogoshi & Takahashi, 2019), language acquisition (Taniguchi, 2019), language teaching research (Tajino, Smith, & Kanamaru, 2019), lifelong learning (Smithers, 2019), Team Teaching (Pearce, 2019), and vocabulary teaching (Sasao, 2019).
3. ALTs are sometimes referred to as AETs (Assistant English Teachers).

References

Allwright, D. (2003). Exploratory practice: Rethinking practitioner research in language teaching. *Language Teaching Research*, *7*, 113–41.

Allwright, D. (2005a). From teaching points to learning opportunities and beyond. *TESOL Quarterly*, *39*(1), 9–31. doi: 10.2307/3588450

Allwright, D. (2005b). Developing principles for practitioner research: The case of exploratory

practice. *The Modern Language Journal*, *89*(3), 353–366. doi: 10.1111/j.1540-4781.2005.00310.x

Allwright, D., & Bailey, K. M. (1991). *Focus on the language classroom: An introduction to classroom research for language teachers*. Cambridge: Cambridge University Press.

Allwright, D., & Hanks, J. (2009). *The developing language learner: An introduction to exploratory practice*. Basingstoke: Palgrave Macmillan.

Atkinson, C., Eldabi, T., & Paul, R. J. (2002). Integrated approaches to health informatics research and development. *Logistics Information Management*, *15*(2), 138–152. doi:10.1108/09576050210414015

Brumby, S., & Wada, M. (1990). *Team teaching.* Harlow, Essex: Longman.

Carless, D. R. (2006). Good practices in team teaching in Japan, South Korea and Hong Kong. *System*, *34*(3), 341–351. doi:10.1016/j.system.2006.02.001

Checkland, P. (1983). OR and the systems movement: Mapping and conflicts. *Journal of the Operational Research Society, 34*(8), 661–675. doi:10.2307/2581700

Checkland, P. (2001). Soft systems methodology in action: Participative creation of an information strategy for an actual hospital. In J. Rosenhead & J. Mingers (Eds.), Rational analysis for a problematic world revisited (pp. 91–113). Chichester, West Sussex: Wiley.

Checkland, P., & Poulter, J. (2006). *Learning for action: A short definitive account of Soft systems methodology and its use for practitioners, teachers and students*. Chichester, West Sussex: John Wiley & Sons.

Checkland, P., & Scholes, J. (1999). *Soft systems methodology in action*. Chichester, West Sussex: John Wiley & Sons.

Connell, N. A. D. (2001). Evaluating soft OR: some reflections on an apparently 'unsuccessful' implementation using a soft systems methodology (SSM) based approach. *Journal of the*

Operational Research Society, *52*(2), 150–160. doi:10.1057/palgrave.jors.2601054

Davison, C. (2006). Collaboration between ESL and content teachers: How do we know when we are doing it right? *International Journal of Bilingual Education and Bilingualism, 9*(4), 454–475. doi:10.2167/beb339.0

Ferrari, M., Fares, B., & Martinelli, P. (2002). The systemic approach of SSM: The case of a Brazilian company. *Systemic Practice and Action Research, 15*(1), 51–66. doi:10.1023/A:1014241704857

Gladman, A. (2015). Team teaching is not just for teachers! Student perspectives on the collaborative classroom. *TESOL Journal, 6*(1), 130–148. doi:10.1002/tesj.144

Hanks, J. (1999). Enthusiasm, puzzlement, and exploratory practice. *International House Journal of Education and Development* (7), 14–16.

Hanks, J. (2009). Inclusivity and collegiality in Exploratory Practice. In T. Yoshida, H. Imai, Y. Nakata, A. Tajino, O. Takeuchi, & K. Tamai (Eds.), *Researching Language Teaching and Learning: An integration of practice and theory* (pp. 33–55). Bern: Peter Lang.

Hanks, J. (2017). *Exploratory practice in language teaching: Puzzling about principles and practices*. London: Palgrave Macmillan.

Hanks, J. (2019a). From research-as-practice to exploratory practice-as-research in language teaching and beyond. *Language Teaching, 52,* 143–187.

Hanks, J. (2019b). Involving everyone in enhancing quality of life in language education: Explorations and insights from praxis. *JACET International Convention Selected Papers*, *6*, 3–28.

Hiratsuka, T. (2019). Assuring the quality of classroom life through exploratory practice: Learner's experiences. *JACET International Convention Selected Papers*, *6*, 101–117.

Holliday, A. (1990). A role for soft systems methodology in ELT projects. *System*, *18*(1), 77–84.

doi:10.1016/0346-251x(90)90030-9

Hosogoshi, K., & Takahashi, S. (2019). A 'softer' instructional systems design for language pedagogy. In A. Tajino (Ed.), *A systems approach to language pedagogy* (pp. 83–98). Singapore: Springer

Johannes, A. A. (2012). Team teaching in Japan from the perspectives of the ALTs, the JTEs, and the students. *TEFLIN Journal*, *23*(2), 165–182. doi:10.15639/teflinjournal.v23i2/165-182

Kato, Y., & Dalsky, D. (2019). Interaction in the language classroom: A systems approach. In A. Tajino (Ed.), *A systems approach to language pedagogy* (pp.121–132). Singapore: Springer.

Liu, L. (2008). Co-teaching between native and non-native English teachers: An exploration of co-teaching models and strategies in the Chinese primary school context. *Reflections on English Language Teaching*, *7*(2), 103–118. Retrieved from http://www.nus.edu.sg/celc/research/books/relt/vol7/no2/103-118liu.pdf

Luo, W. H. (2010). Collaborative teaching of EFL by native and non-native English-speaking teachers in Taiwan. In A. Mahboob (Ed.), *The NNEST lens: Non native English speakers in TESOL* (pp. 263–284). Newcastle upon Tyne: Cambridge Scholars Publishing.

Maswana, S., & Tajino, A. (2020). ESP curriculum development: A systems approach. In H. Terauchi, J. Noguchi, & A. Tajino (Eds.), *Towards a new paradigm for English language teaching* (pp. 66–76). Abingdon: Routledge.

McConnell, D. L. (2000). *Importing diversity: Inside Japan's JET program.* California: University of California Press.

Miyazato, K. (2009). Power-sharing between NS and NNS teachers: Linguistically powerful AETs vs. culturally powerful JTEs. *JALT Journal*, *31*(1), 35–62. Retrieved from http://jalt-publications.org/jj/articles/271-power-sharing- between-ns-and-nns-teachers-linguistically-powerful-aets-vs-culturally

Ng, M. L. (2015). Difficulties with team teaching in Hong Kong kindergartens. *ELT Journal, 69*(2), 188–197. doi:10.1093/elt/ccu057

Park, J. E. (2014). English co-teaching and teacher collaboration: A micro-interactional perspective. *System*, *44*, 34–44. doi:10.1016/j.system.2014.02.003

Pearce, D. R. (2019). Team teaching and team learning: SSM applied to the team-taught classroom. In A. Tajino (Ed.), *A systems approach to language pedagogy* (pp.99–119). Singapore: Springer.

Pearce, D.R., Stewart, T., & Tajino, A. (2019). Making a difference through quality co-teaching between native and non-native teachers in foreign language classrooms. In J. D. D. M. Agudo (Ed.), *Quality in TESOL and teacher education: From a results culture towards a quality culture* (pp.103-112). Abingdon, Oxon.:Routledge.

Perry, B., & Stewart, T. (2005). Insights into effective partnership in interdisciplinary Team Teaching. *System, 33*(4), 563–573. doi:10.1016/j.system.2005.01.006

Rowland, L. (2011). Lessons about learning: Comparing learner experience with language research. *Language Teaching Research*, *15*, 254–267. doi: 10.1177/1362168810388726

Sakamoto, N. (2020). *A study on identity transformation through narratives of native assistant language teachers in Japanese public secondary schools* (Unpublished doctoral dissertation). The University of Hyogo, Hyogo Prefecture.

Sasao, Y. (2019). Vocabulary teaching: A systemic perspective. In A. Tajino (Ed.), *A systems approach to language pedagogy* (pp.39–47). Singapore: Springer.

Slimani-Rolls, A., & Kiely, R. (2014). 'We are the change that we seek': Developing teachers' understanding of their classroom practice. *Innovations in Education and Teaching International*, *51*, 425–435. doi: 10.1080/14703297.2014.894328

Smith, C. (2009). Developing a new English for academic purposes course: Administrator- teacher-

student collaboration. In T. Yoshida, H. Imai, Y. Nakata, A. Tajino, O. Takeuchi, & K. Tamai (Eds.), *Researching language teaching and learning: An integration of practice and theory* (pp. 95–113). Bern: Peter Lang.

Smithers, R. W. (2019). A soft systems approach to lifelong learning. In A. Tajino (Ed.), *A systems approach to language pedagogy* (pp. 69–81). Singapore: Springer.

Stewart, T. (2018). Team teaching collaborations: Contact, conflict, and empowerment. *JACET Journal, 62,* 29-47.

Stewart, T. (2019). Surveying the landscape of second language classroom research. In A. Tajino (Ed.), *A systems approach to language pedagogy* (pp.49–67). Singapore: Springer.

Stewart, T., Dalsky, D., & Tajino, A. (2019). Team learning potential in TESOL practice. *TESOL Journal*, *10*(3), 1–12. doi: 10.1002/tesj.426

Tajino, A. (2002). Transformation process models: A systemic approach to problematic Team Teaching situations. *Prospect, 17*(3), 29–44. Retrieved from http://www.ameprc.mq.edu.au/docs/prospect_journal/volume_17_no_3/17_3_3_Tajino.pdf

Tajino, A. (2009). Understanding life in the language classroom: A systemic approach. In T. Yoshida, H. Imai, Y. Nakata, A. Tajino, O. Takeuchi, & K. Tamai (Eds.), *Researching language teaching and learning: An integration of practice and theory* (pp. 115–130). Bern: Peter Lang.

Tajino, A. (2019). Teaching English grammar for communicative purposes: A systems approach. In A. Tajino (Ed.), *A systems approach to language pedagogy* (pp.23–38). Singapore: Springer.

Tajino, A., James, R., & Kijima, K. (2005). Beyond needs analysis: Soft systems methodology for meaningful collaboration in EAP course design. *Journal of English for Academic Purposes*, *4*(1), 27–42. doi: 10.1016/j.jeap.2004.01.001

Tajino, A., & Smith, C. (2016). Beyond team teaching: An introduction to team learning in language education. In A. Tajino, T. Stewart & D. Dalsky (Eds.), *Team teaching and team learning in the language classroom: Collaboration of innovation in ELT* (pp. 11–27). Abingdon, Oxon.: Routledge.

Tajino, A., Smith, C., & Kanamaru, T. (2019). Toward managing language pedagogy and language teaching research. In A. Tajino (Ed.), *A systems approach to language pedagogy* (pp.1–10). Singapore: Springer.

Tajino, A., Stewart, T., & Dalsky, D. (Eds.). (2016). *Team teaching and team learning in the language classroom: Collaboration for innovation in ELT.* Abingdon, Oxon.: Routledge.

Tajino, A., & Tajino, Y. (2000). Native and non-native: What can they offer? Lessons from team teaching in Japan. *ELT Journal*, *54*(1), 3–11. doi:10.1093/elt/54.1.3

Tajino, A., & Walker, L. (1998). Perspectives on team teaching by students and teachers: Exploring foundations for team-learning. *Language, Culture and Curriculum, 11*(1), 113–131. doi:10.1080/07908319808666544

Taniguchi, K. (2019). Language acquisition: A systemic view from cognitive linguistics. In A. Tajino (Ed.), *A systems approach to language pedagogy* (pp. 11–21). Singapore: Springer.

Terauchi, H., & Maswana, S. (2019). Systems thinking: An ESP genre approach. In A. Tajino (Ed.), *A systems approach to language pedagogy* (pp. 133–146). Singapore: Springer.

Translanguaging and Trans-semiotizing Approaches to Content and Language Integrated Learning (CLIL): Innovating With the Multimodalities-Entextualization Cycle (MEC)

Peichang (Emily) He, Faculty of Education, The University of Hong Kong, Hong Kong

Angel M. Y. Lin, Faculty of Education, Simon Fraser University, Canada

Abstract

This research explores how translanguaging (Garcí & Li, 2014) and trans-semiotizing, approaches can facilitate students' expansion of their communicative repertoires to gradually make new ways of speaking, writing, and unfamiliar registers their own (Lin, 2012, 2017, 2019; He, Lai & Lin, 2017). Drawing on the sociocultural and social-semiotic perspectives, the study investigates meaning making in a multilingual CLIL classroom through fine-grained analysis of a series of lesson observation videos, teaching materials and sample student works collected during a school-university collaborative project on building up the CLIL capacity of teachers and South Asian minoritized students in Hong Kong. The results of this study shed light on how teachers design spaces for translanguaging and trans-semiotizing (Lin, 2019; Lin, Wu & Lemke, 2020) and spaces for target language and register use in the different stages of a curriculum genre (Rothery, 1996). The Multimodalities-Extextualization Cycle (MEC) is proposed and discussed as a heuristic curriculum genre for CLIL education (Lin 2015b, 2016). The study concludes by discussing theoretical and pedagogical implications and providing implementation suggestions for CLIL curriculum and teacher professional development (He & Lin, 2018; Lo, 2020).

Keywords: translanguaging, trans-semiotizing, multimodal approaches, Content and Language

Integrated Learning (CLIL), Multimodalities-Entextualization Cycle (MEC)

Introduction

Content and Language Integrated Learning (CLIL) is a rapidly growing area of both research and practice in many parts of the world, especially in Europe and Asia where an increasing number of schools and universities are using CLIL as an educational approach to integrate the learning of content with the learning of a second, foreign, or additional language (Coyle, Hood, & Marsh, 2010). However, due to the special characteristics of CLIL and various contextual complexities, it remains a challenge for teachers to achieve the integration of content and language teaching in CLIL classrooms (Dalton-Puffer, 2018; Lin, 2016; Lo, 2020; Ruiz de Zarobe, 2016). One of the recurring themes for connecting content and academic language for English learners is to develop academic language across all curriculum areas through developing opportunities for a curriculum of talk (Swinney & Velasco, 2011). But how do we help students to make 'alien' words and foreign ways of speaking/writing/thinking in foreign/additional languages/academic registers their own? Lemke (1990) proposed that teachers should help students to grasp semantic and conceptual relationships in colloquial language first and then guide them to substitute scientific and technical terms for colloquial words. Recent research on co-development of science literacy and academic language literacy indicated that both designed scaffolding and spontaneous scaffolding (Gibbons, 2009; Lin, 2016) are indispensable for the development of content knowledge and academic language knowledge (He & Lin, 2019). Empirical studies on CLIL also pointed out that translanguaging (Garcí & Li, 2014) and trans-semiotizing (Lin, 2015a) can be useful scaffolding for meaning making and the development of content and academic language (Karlsson, Nygård Larsson & Jakobsson, 2019; Lin, 2019; Lin & He, 2017; Lin & Wu, 2015).

In this study, we shall focus on the CLIL education of multilingual and multicultural students

in an English-Medium-Instruction (EMI) science classroom of a Hong Kong secondary school. We shall draw on recent theorization of translanguaging (Lin, Wu & Lemke, 2019) and multimodal approaches (Danielsson, 2016) to examine how translanguaging and multimodal activities can facilitate ethnic minority students' expansion of their communicative repertoires (Lin, 2012) to gradually make foreign ways of speaking/writing and unfamiliar registers their own. We shall also adopt the Multimodalities-Extextualization Cycle (MEC) (Lin 2015b, 2016) as a heuristic tool to analyse how the CLIL teacher can design spaces for translanguaging and trans-semiotizing and spaces for target language/register use in the different stages of a curriculum genre (Rothery, 1996) to achieve the objectives of both content and language development in the CLIL classroom.

Literature Review

Meaning Making Through Bridging Between Colloquial and Academic Registers

From social semiotic perspectives, human learning is a meaning making process—a semiotic process during which language is the primary semiotic resource to construe (i.e. to construct and understand) content (Halliday, 1993). Language is essential for knowledge construction and interpretation in learning content subjects, and is viewed as the principal means through which the conceptual meaning of academic subjects is co-constructed in the classroom (Lemke, 1988). Academic subjects are characterized by their technical language features that are alien to learners. Content knowledge will be incomprehensible if the "foreign" language is not "translated" into students' own language; namely, the colloquial language which they use for daily communication. However, as academic literacy is a stipulated syllabus outcome and the only legitimate discourse in school assessments, students need to develop academic literacy for entry to higher education institutions. Therefore, to facilitate academic subject learning, teachers need to help students to bridge the gap between colloquial language and academic language and guide them to "translate back

and forth" between the two different registers, as emphasized by Lemke (1990) in science education, "Students understand best what is explained to them in the language they use themselves, ordinary colloquial English... Students will begin to grasp semantic and conceptual relations in colloquial language first. Then they will substitute scientific, technical terms for colloquial words" (p. 172-173). It is worth noting that, in the teaching of abstract and technical subjects such as science, the bridging of languages may go beyond lexico-grammatical features and extend to the hybridizing of different registers. As proposed by Lemke (1990), "Teachers should use all the stylistic and rhetorical means available to communicate science to students, including narrative and dramatic presentations; humor, irony, and metaphor; fiction and fantasy; reference to actual scientific activities, disputes, and persons; personal anecdotes and historical examples" (p.174). Such trans-registering practices help to broaden the range of designed scaffolding and spontaneous scaffolding strategies (Gibbons, 2009; Lin, 2016) which allow teachers to deploy flexibly semiotic resources from an expanded communicative repertoire (Lin, 2012) to help students to bridge the gap between colloquial and academic language in content-based instruction.

Translanguaging and Multimodal Amplification of Meaning

Although language is the primary semiotic resource in human learning, meaning making is never the result of language alone, but the interplay between language (i.e. verbal) and other meaning making resources such as visual and gestural semiotics during communication. According to Danielsson, (2016), each mode of presentation has its specific "affordance" for meaning making defined as "meaning potential" (van Leeuwen, 2005). Through strategic deployment of multimodal resources, the different semiotic modes interact simultaneously, backgrounding and foregrounding each other so that the meaning of concepts is not just repeated but also amplified (Kress, Jewitt, Ogborn &Tsatsarelis, 2001). As Lemke (1998) elucidated, "The 'concepts' of science are not verbal

concepts...They are semiotic *hybrids,* simultaneously and essentially verbal-typological and mathematical-graphical-operational-topological. The actional, conversational, and written textual genres of science are historically and presently, fundamentally and irreducibly *multimedia genres*" (p.87, emphasis in original). The multimodal approaches to meaning making have proved to be fundamental and fruitful in recent studies in science education (Cheng & Gilbert, 2015; Danielsson, 2016).

The multimodal social semiotic views of teaching and learning is essential for CLIL classroom where non-language content subjects are taught in a foreign/second/additional language. In bi/multilingual education, translanguaging has been widely researched since its emergence as a pedagogy that allows bilingual learners to comprehend the subject matter in one language and express their ideas about the subject in another (Baker, 2011). Translanguaging as "*multiple discursive practices* in which bilinguals engage in order to *make sense of their bilingual worlds*" (Garcí, 2009, p.45, emphasis in original) has been found to be beneficial for the content and language development of emergent bilinguals (Garcí, 2009) through promoting their understanding of the subject matter and learning of the less familiar language (i.e., the foreign/additional language) (Baker, 2011; Lin & He, 2017; Lin & Wu, 2015). It should be noted that, the translanguaging approach questions the concept of language separation and rejects the view that assumes bilingualism as the addition of two separate languages or the practice of parallel monolingualisms (Garcí & Lin, 2017; Lin & Lo, 2018); rather, it advocates a holistic view of semiotic (i.e. meaning-making) repertoire which "signals a trans-semiotic system with many meaning-making signs, primarily linguistic ones that combine to make up a person's semiotic repertoire" (Garcí & Li, 2014, p.42). Building on Halliday's (2013) "trans-semiotic" view, Lin (2015b) developed the concept of "trans-semiotizing" and emphasised the seamless flow of entanglement of multiple meaning making resources. Languages (as a central semiotic) not only interact with each other but also intertwine with other

semiotics (e.g., visual images and gestures) in human communication practices during which the common semiotic repertoire expands with the ongoing contributions of the communicators (Lin, 2012; Lin, 2019). This has been evidenced in a study investigating the effective communication between a professor and a multilingual audience through translanguaging and trans-semiotizing during a multimodal seminar presentation, despite the fact that the speaker and the audience shared little common language background (He, Lai, & Lin, 2017).

Recent literature illustrates the multilingual, multimodal, experiential and transformative nature of translanguaging. In a study of how transnational self-directed learners of Chinese mobilized their multilingual, multimodal, and multisemiotic repertoires, as well as their learning and work experiences as resources in language learning, the researchers witnessed how this translanguaging and multimodal learning process transcended the limits of form, meaning and function in traditional language learning and created a "space" for the multilingual learners to bring together and orchestrate their personal beliefs, histories, experiences, physical and cognitive abilities, social ideology as well as the environment, hence transforming language learning into a multilingual, multimodal and multisensory experience (Li & Ho, 2018). Translanguaging is theorized as both a practice during which participants are engaged in "dynamic" and "functionally integrated" uses of languages or language varieties, and a process of knowledge construction that goes beyond language systems (Li, 2018). Translanguaging classroom discourses not only encourage fluid multilingual practices within the limits set up by the roles and objectives of tasks but also aim to push the limits and transcend the boundaries, hence transforming the classroom experiences (Li & Lin, 2019). This dynamic, fluid and transformative nature of translanguaging resonates with the dynamic and distributed view of translanguaging as *flows* (Lemke, 2016, 2018; Lin, Wu & Lemke, 2019). According to Lemke (2018), translanguaging has a material history, which consists of processes and *flows* across timescales and complex eco-social systems. Each individual body is a place where the *flows* of the

community, past and present, have passed through it. Translanguaging is a community phenomenon within which individuals as mediums for the *flows* of matter, energy and information are interconnected with those of other members in the community, and the *flows* in different communities are further intertwined in the ecology, which makes translanguaging a dynamic, fluid and constantly intertwining ecosystem phenomenon.

The Multimodalities-Extextualization Cycle (MEC) as a curriculum genre for CLIL

As discussed above, the translanguaging and multimodal approaches are innovative for CLIL classroom practices. In this study, we shall discuss how teachers can design spaces for translanguaging and trans-semiotizing and spaces for target language/register use in the different stages of a curriculum genre (Rothery, 1996). The MEC (Lin 2015b, 2016) has been proposed as a heuristic curriculum genre to achieve this aim. The MEC has three stages (Figure 1): first, create a rich experiential context; second, engage students in reading and note-making; and third, engage students in entextualizing the experience.

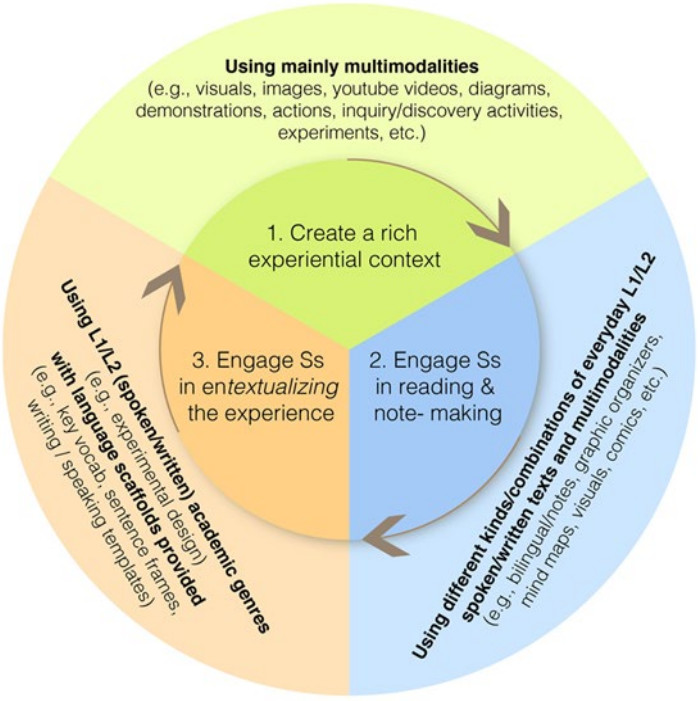


Figure 1. The Multimodalities-Entextualization Cycle (MEC) (Key: Ss = students) .

(Adapted from Lin, 2015b)

In the MEC, while Stages 1 and 2 allow for the uninterrupted flow of meaning-making and pedagogical support through translanguaging and trans-semiotizing, the third stage allows students to have a space to practise, orienting their meaning making towards the discourse and cultural patterns required by the school and academia for successful participation in future assessment tasks and for expanding their communicative repertoires (Lin, 2012). In this stage, scaffolding needs to be provided (e.g. useful vocabulary, sentence patterns, writing/speaking starters). The MEC in principle can be reiterated without an end-point to emphasize the equal importance of all the multiple linguistic and multimodal resources. Hence, the MEC is proposed as an innovative tool for CLIL educators to think about how to design systematic scaffolding in content-based language education and CLIL classrooms.

Method

This research was part of a school-university partnership QEF Project[i] funded by the Education Bureau of the Hong Kong government. The research project aimed at building up the CLIL capacity of teachers and addressing learner diversity in multicultural and multilingual EMI science classrooms. The project lasted for 18 months and the researchers provided CLIL curriculum development support for the teachers who taught ethnic minority students of South Asian community backgrounds in the school. Ethical approval by school principal, teachers, students and parents had been obtained with informed consent forms signed before the project started. The current research report focuses on the data collected in the second half of the project and aims to address two research questions:

1. How did the teacher and students make meaning through translanguaging and multimodal activities in the CLIL science classroom?

2. How did the teacher create spaces of translanguaging and trans-semiotizing and spaces for academic language development through a MEC-based curriculum genre?

Research site and participants

Due to teacher and learner diversity, the collaborative school was characteristised by its multilcultural and multilingual background. It enrolled both the local Cantonese-speaking students and students from ethnic minority families from South Asian countries; e.g., Pakistan, India and Nepal. The Cantonese-speaking students were streamed into Chinese-Medium-Instruction (CMI) classes while the South Asian students were streamed into EMI ones. The teaching staff consisted of local bilingual Cantonese-English speaking teachers, native English-speaking teachers and South Asian teachers who spoke English and Urdu, Hindi or Nepalese. Although both English and Cantonese were the working languages at the school, the languages of the South Asian communities

could also be heard especially in the multilingual EMI classes. As the Hong Kong government implements the "biliteray and tri-lingualism" language policy, a *Putonghua* course was also available in the schedule which not only provided an opportunity for students to learn a new spoken language, but also allowed the South Asian students to learn some traditional Chinese characters (繁體字) which constitute one of the official written languages in Hong Kong.

Ms Wan (pseudonym) was a Cantonese-speaking teacher with three years' teaching experience. She taught biology, integrated science (IS) and integrated humanity subjects for both CMI and EMI classes in the school. The teacher was keen on improving her students' English academic literacy. During the collaboration, she was studying in a CLIL MEd Programme in a local university. Ms Wan was fluent in both Cantonese and English, but she did not know the South Asian students' first languages. She had developed a friendly relationship with the students and tried to learn some basic Urdu words from the students so that she could use them to attract the students' attention during the lessons. The student participants were in a S3 (Secondary 3 / Year 9) EMI class consisting of 14 girls from Pakistani and Indian families. Although the students all spoke Urdu, few of them knew how to read or write the language. Their proficiency of English and Cantonese showed a huge diversity. Some students had relatively better English proficiency as they went to EMI primary schools. Others were born in Hong Kong and had acquired some Cantonese through daily exposure to the local Cantonese community; however, their English proficiency could be weak as not all of them studied in an EMI primary school and the English proficiency of their parents was generally low. There were still some students who had just immigrated to Hong Kong and knew little English, Chinese or Cantonese. The linguistic diversity had increased the teaching and learning challenges in the CLIL classes.

Procedure and Materials

During the project, the researcher (first author) was stationed at the collaborative school as a teaching consultant and co-developed CLIL materials with the participating teachers for the EMI science classes. The data for this article were based on the collaboration with Ms Wan in her S3 IS lessons. During the collaboration, the researcher provided Ms Wan with the lesson analysis and teaching resources of the unit to be taught. Ms Wan was able to understand the content and language objectives of the lesson design, and would provide feedback on the teaching materials. She would also share with the researcher her own plans about how to incorporate the materials into the lessons so that the CLIL lessons were implemented based on the materials and pedagogy co-developed by Ms Wan and the researcher.

The data collection was carried out when Ms Wan was teaching the circulatory system unit. The teaching of the whole unit lasted for six weeks during which the researcher observed and took videos of the classes including four lessons in the classroom and two experiments in the lab. As the teacher and the researcher had tried out their collaboration in the previous unit, both Ms Wan and the students had got used to the co-developed curriculum and lesson observation by the researcher, and they did not pay much attention to the camera during the lessons. The data collected in this study included three sources: first, six lesson observation videos, each recording a 70-minute double-period lesson; second, the teaching and learning materials as well as samples of student work; and third, a 92-minute semi-structured interview with Ms Wan and a 53-minute focus group interview with the students at the end of the project. The data collected were analysed in two rounds: First, initial data analysis was made following an inductive coding method (Miles & Huberman, 1994); the unclear scenarios were identified and then clarified by the teacher and students during the follow-up interviews; for example, the Urdu phrases mentioned by Ms Wan were explained by the teacher and translated by the students. Second, the researchers conducted fine-grained multimodal analysis of the data collected based on both the translanguaging and multimodal approaches to meaning making

(Danielsson, 2016; Lemke, 2016; Lin, Wu & Lemke, 2019) and the MEC (Lin, 2015b; 2016), with foci on the processes of translanguaging (Garcí & Li, 2014), trans-semiotizing(Lin, 2015a) and trans-registering (Lin, Wu & Lemke, 2019) during teacher-student interactions at different stages of the MEC.

Results

In this section, we shall introduce the results of the research by first presenting a detailed analysis of the beginning sections of the first lesson of the collaborative unit to illustrate the translanguaging and multimodal meaning making through interactions between Ms Wan and the students. We shall then analyse the revision lesson to illustrate how spaces for academic literacy development were created in the CLIL classroom based on the MEC (Lin, 2015b; 2016).

In the first lesson on the circulatory system, Ms Wan intended to make an overall introduction about the key components of the unit, and then went on to teach the first component “blood”. As shown in Figure 2, the teacher prepared a worksheet which consisted of the relevant concepts and interrelationships to be taught in the early stages of the first lesson.

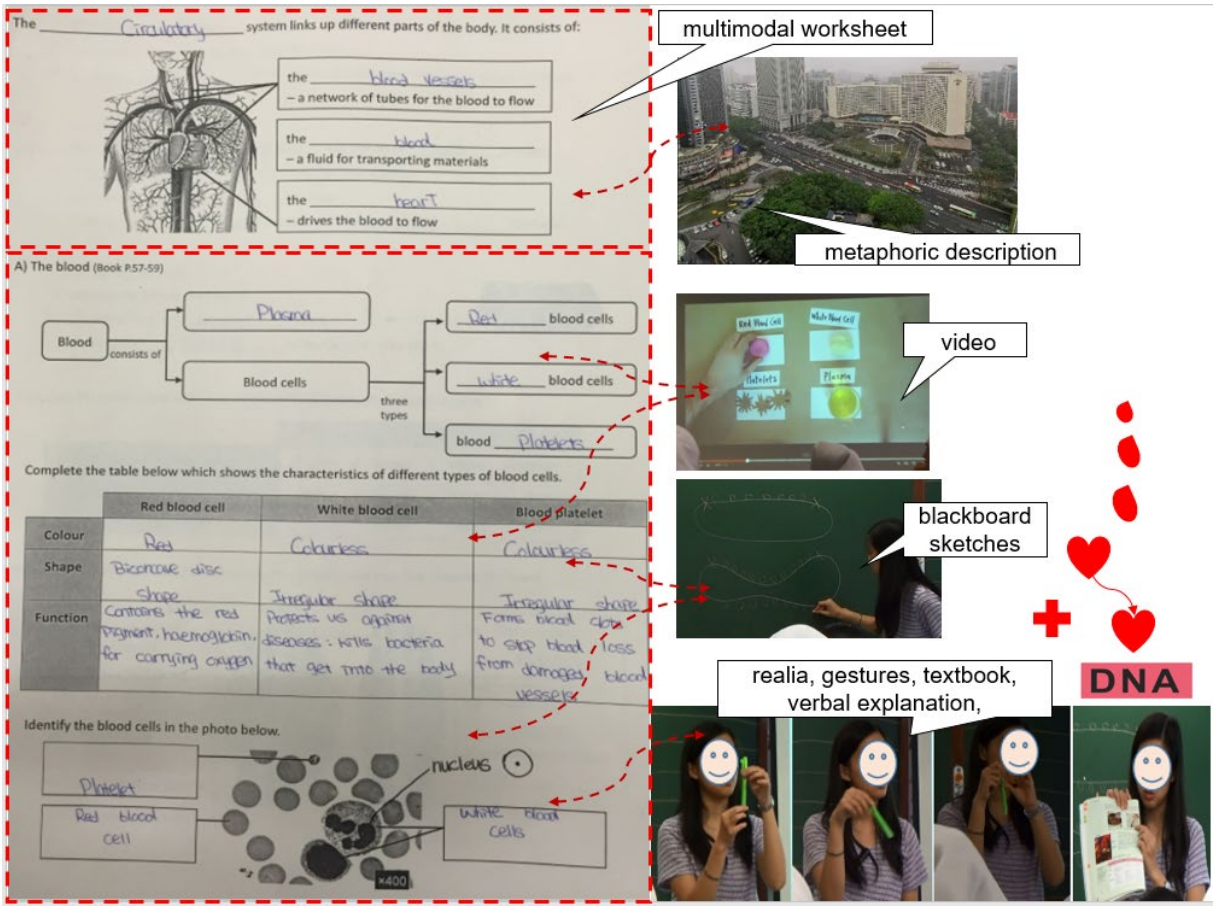


Figure 2. Translanguaging and multimodal activities in the first lesson of the circulatory system unit.

Introduction to the circulatory system

The circulatory system unit was composed of three main sections: "blood", "blood vessels" and "heart". As the technical term "circulatory system" may sound unfamiliar to the students, Ms Wan provided a diagram on the worksheet (Figure 3) with which she intended to help the students to build up the schemata of the unit through guiding them to identify the components in the diagram and label them accordingly.

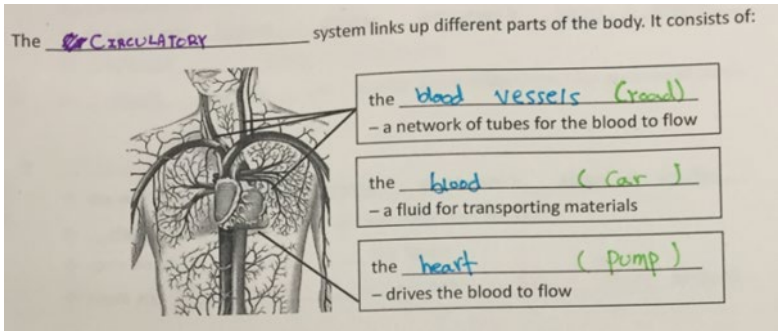


Figure 3. Diagram of the circulatory system on a student's worksheet.

Ms Wan started the lesson by writing the title "Circulatory System" on the blackboard. At the beginning of the lesson, as the students were very excited talking in Urdu and English, Ms Wan tried to calm them down by uttering a few sounds like "Choop" and "Shuno" which she learned from the students, meaning "Silence" and "Listen to me" respectively in their Urdu language. Since the key word "circulatory" is a technical term which is unfamiliar to the students, Ms Wan reminded them of its location in the textbook and also tried to paraphrase the root word "circulate" by an everyday expression "going round and round", with an accompanying gesture by her index finger drawing circles in the air, which a student further interpreted as "in a circle". Although students might have heard about "blood", "blood vessel" and "heart", they might not understand the scientific meanings of the terms. Ms Wan thus used city transportation as a metaphor to explain the features and functions of the circulatory system.

Excerpt 1: "Yes, you can imagine it, like in a city."

T: Today we'll start a new system.

Ss: Circulatory system.

[*The teacher wrote the topic "Circulatory System" on the blackboard and then started booting the computer. The students were chatting excitedly in both English and Urdu, their first language (L1).*]

T: All right. **Choo::p**. [*A word sounds like the Urdu word meaning "Silence"*] So, for circulatory system, this word

itself means circulation. Circulation means, all right, page 57. So "circulate" means, things go...[*talking with her right index finger drawing circles in the air*]

[*Students kept on chatting in both English and their L1 which made the classroom noisy.*]

T: **Shu::no.** [*The teacher changed to an Urdu word (meaning "Listen") to attract students' attention*.] Well, you talk and I talk. We can't listen to each other. All right, it means, things are going round and round.

S1: In a circle.

T: Yeah. [*Teacher distributed the worksheets to students*]

T: There are three parts in our circulatory system. So we need to bring things round and round

S1: Round and round?

T: Yes. So the first thing is *like*, if you got a city, you want people to go round the city, so at least you need some roads.

S2: Circulatory road.

T: [*Turning to the blackboard and started to write*] Yeah. For us, in our body we don't have those kinds of circulatory roads. We don't call it roads.

S3: We have the vessels.

T: Yeah we call it "vessels" [*writing "blood vessels" on the blackboard while speaking*]

S4: Vessels.

T: All right. Which looks like the roads [*Adding "(roads)" on the blackboard after "blood vessels"*] Okay. Even though you got roads, people may not be able to go round. They still need cars, buses...

S2: Blood.

T: [*Went on writing on the blackboard "blood (cars)"*] Yes. In our circulatory system, it means the blood, which is like the cars.

S5: Public cars or private.

T: Yes. It can be public, it can be private.

S2: How do we know these are public cars?

T: Hmm, in our body, we got different types of blood cells. Just like we got different kinds of cars in a road.

S3: Red blood cells and white blood cells.

S4: Mercedes, Ferrari.

S6: BMW.

T: Well I mean bus and mini bus.

S7: Lamborghini.

T: All right. We're talking about the blood. And the last one... [*Turning to the blackboard and writing "heart'*] is the heart. So this kind of cars in your body are very...

S4: Wheels.

T: Lazy. They don't go by themselves.

S3: They just...

T: They need a pump [*went on writing "(pump)" behind "heart"*] to push them round.

S4: Petrol, petrol.

T: Kind of. Yeah.

S2: Traffic light.

T: A pump that gives it energy to go. Okay? So today we'll start with the blood first. We'll start with the blood first, and then blood vessels.

S8: [*Pointing at the blackboard writing*] So the cars are like the blood in the circulatory system?

T: [*Pointing to the words on the blackboard*] Yes, you can imagine it, like in a city.

The metaphoric example seemed to have effectively engaged the students who started to join in the teacher's narrative about the components of the circulatory system by curiously guessing and raising questions following the teacher's talk. The students were able to label the concepts in the

diagram correctly. Some of them (Figure 3) even copied the metaphors “car”, “road” and “pump” beside the corresponding concepts to remind themselves of the characteristics of the key components in the circulatory system.

Blood

The second part of the lesson was about “blood”. The knowledge points included the composition of blood, the characteristics of different types of blood cells, and the identification of blood cells according to their features. Ms Wan designed multimodal exercises with a graphic organizer, a table and a microphotograph to represent the concepts and their inter-connections. To visualize the abstract concepts, she played a video about “the components of blood”. As the presenter in the video demonstrated the four components of blood -- red blood cells, white blood cells, platelets and plasma, by using red play-doh, white play-doh, irregularly shaped cardboard pieces and a glass of yellowish water respectively, the vivid representations helped students to complete the graphic organizer easily. Many students not only fill in the blanks with words but also use color pens to note and draw beside the concepts the corresponding features they captured from the video description (Figure 4).

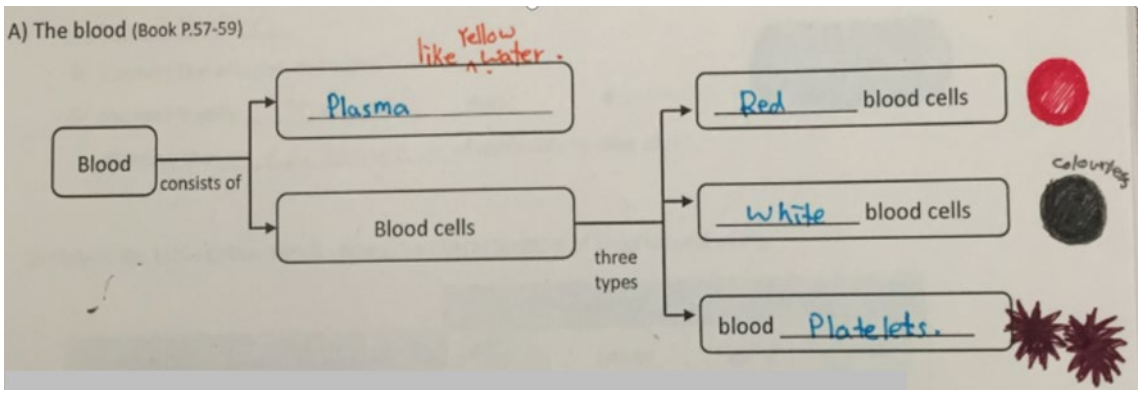


Figure 4. Graphic organizer sorting out components of blood on a student’s worksheet.

However, although the video provided some relevant information about the lesson, during the worksheet exercises, Ms Wan needed to help students to clarify their understanding of some concepts, guided them to think about the causes of some phenomena, and answered students' questions that are relevant to the theme--components of blood. For example, Ms Wan reminded students that using white play-doh as an example to illustrate white blood cells in the video could be misleading as white blood cells are not "white" but "colourless". Similarly, when the students quoted the metaphor "donut" from the video to describe the shape of a red blood cell, Ms Wen suggested a more scientific technical term "biconcave" and she elaborated on the shape by drawing a cross-sectional view of biconcave on a worksheet projected on the screen. To help the students understand the technical word, she reminded students of the prefix, "bi- means two, like bicycle". She then used both her hands to make a quick "going down" action, and with the gesture she translated the term into a colloquial Cantonese word "凹咗"; then pointing to her drawing on the worksheet she added a description, "That means, in the middle part, it's thinner." To guide the students to explore the function of this special characteristic of the shape, Ms Wan raised a question, "Why does it need to be thinner?"

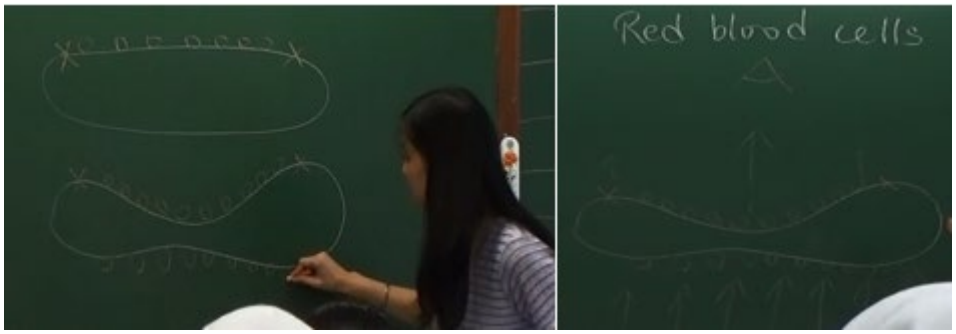


Figure 5. Blackboard sketches about the biconcave shape of a red blood cell.

The teacher then drew sketches of a flat oval-shape and a biconcave on the blackboard (Figure

5). By guiding students to compare the surface areas of the two shapes, Ms Wan explained to them that more oxygen can be carried along if red blood cells have a larger surface area--a biconcave shape. These blackboard sketches of the biconcave shape were once again referred to after several minutes when the teacher and the students discussed the identification of different types of blood cells in a microphotograph. Ms Wan drew an "eye" above the "biconcave (red blood cell)" and some arrows below the shape to show the direction of light in an electron microscope. She told students that due to the thinner of the middle, more lights can go through this part so the observer may find the red blood cell lighter in the centre while darker in the parts around. In a student's worksheet (Figure 6), the teacher's explanation about the special features and functions of the biconcave shape of red blood cell was noted down at the margin of the worksheet with arrows linking the notes with the corresponding concepts in the table or the photo. The student also copied the Chinese translation of the "biconcave disc" (雙凹圓盆) and "blood clot" (血凝塊) from the textbook to amplify the meaning of the concepts to facilitate comprehension and memorization.

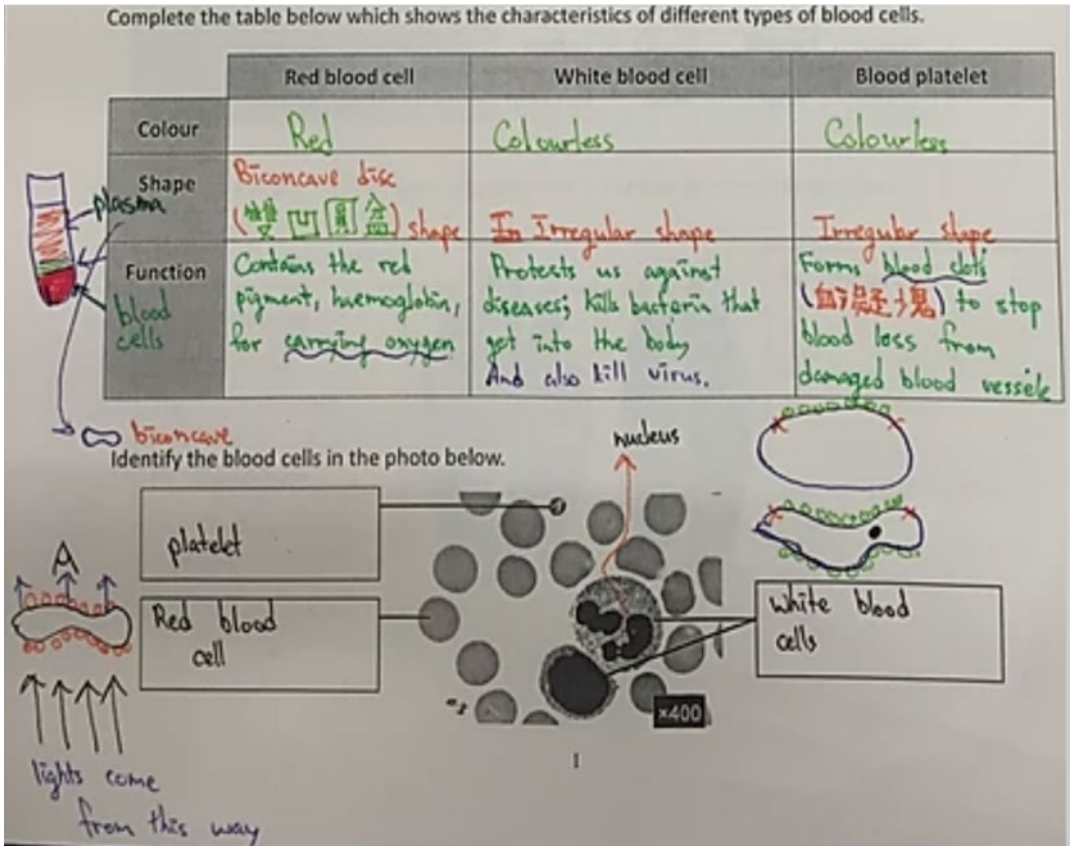


Figure 6. Information in a table and a microphotograph about blood cells on a student's worksheet.

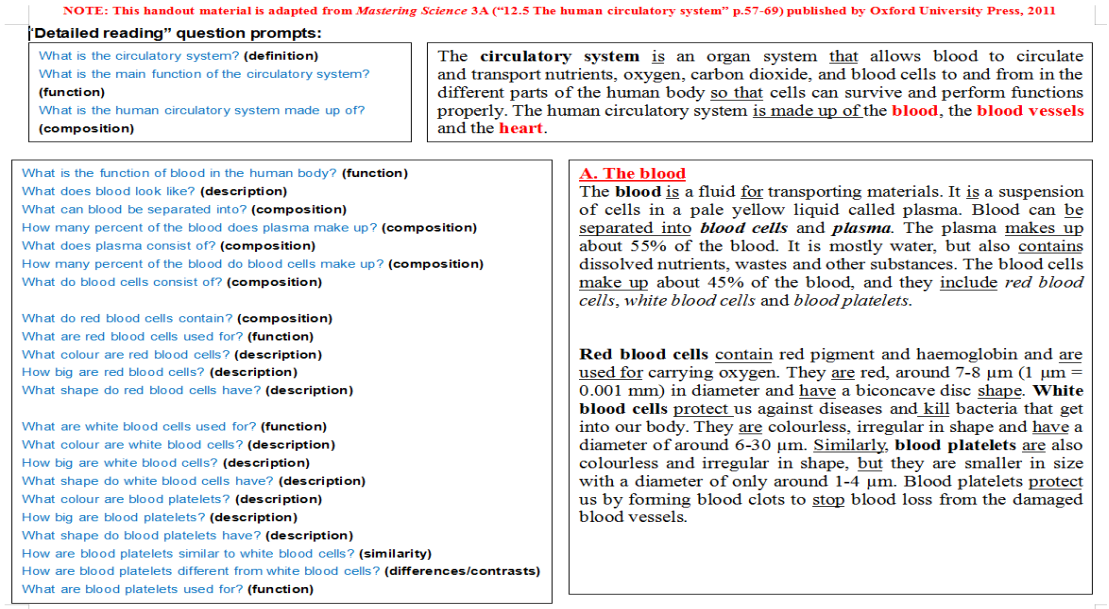
During the lesson, students were intrigued by the different scientific phenomena being discussed and they tried to raise different questions even though some of them could not even complete the sentence of a question due to the lack of corresponding vocabulary in either English or Cantonese. For example, one student seemed to be curious about those people who have problem in blood clotting. She tried to ask the question but could not finish her sentence after several attempts. This confused Ms Wan as she did not understand what the student wanted to know. At this moment, the student and her peers resorted to their first language and gradually co-constructed the question. One of the students whose English proficiency was higher helped her to rephrase the question in English so that Ms Wan was able to understand what they wanted to know and explained the phenomenon. Another student was interested in how blood can be separated into different types of blood cells. To explain the phenomenon, Ms Wan elaborated on the separation of blood cells from the plasma by centrifugation. She used a pen to represent a test-tube with blood and then swirled the "test-tube with blood" in the air quickly to illustrate the rapid spinning of the "test-tube" in a centrifuge. Then she pointed to the upper and lower parts of the "test-tube" to show the separated plasma and blood cells respectively. After the demonstration, Ms Wen showed the students the pictures in the textbook about centrifugation so that they may have clearer images of the centrifuge and centrifuged blood sample (Figure 7).

Figure 7. Demonstration of centrifugation of blood cells by gestures and textbook images.

The discussion on blood and blood cells further reminded the students of other relevant topics such as blood transfusion at the Red Cross, the previous blood donation of one student in the class, Ms Wan's practicum in a lab, and an interesting question "whether DNA will be affected if one receives a blood transfusion?"--the same question which Ms Wan had asked her teacher when she was in secondary school. The lesson thus went on with interweaving of discussion about the worksheet exercises and the teacher's answers to the questions raised by students spontaneously.

Revision

After four lessons of knowledge construction through translanguaging and trans-semiotizing as well as observations and hands-on practices in experiments, Ms Wan decided to draw students' attention to the development of academic language use. However, since the science textbook only consisted of truncated passages that are fragmented with diagrams, tables, activities, and lab reports scattered here and there, teachers seldom referred to the texts and students did not want to read them. After discussing with the researcher, Ms Wan decided to adapt the passages in the textbook and developed a set of genre-based (Rose & Martin, 2012) reading handouts. Figure 8 shows a section of the handouts about the introduction of the circulatory system and the components of "blood".



NOTE: This handout material is adapted from *Mastering Science* 3A ("12.5 The human circulatory system" p.57-69) published by Oxford University Press, 2011

"Detailed reading" question prompts:

What is the circulatory system? **(definition)**
What is the main function of the circulatory system? **(function)**
What is the human circulatory system made up of? **(composition)**

The **circulatory system** is an organ system that allows blood to circulate and transport nutrients, oxygen, carbon dioxide, and blood cells to and from in the different parts of the human body so that cells can survive and perform functions properly. The human circulatory system is made up of the **blood**, the **blood vessels** and the **heart**.

What is the function of blood in the human body? **(function)**
What does blood look like? **(description)**
What can blood be separated into? **(composition)**
How many percent of the blood does plasma make up? **(composition)**
What does plasma consist of? **(composition)**
How many percent of the blood do blood cells make up? **(composition)**
What do blood cells consist of? **(composition)**

What do red blood cells contain? **(composition)**
What are red blood cells used for? **(function)**
What colour are red blood cells? **(description)**
How big are red blood cells? **(description)**
What shape do red blood cells have? **(description)**

What are white blood cells used for? **(function)**
What colour are white blood cells? **(description)**
How big are white blood cells? **(description)**
What shape do white blood cells have? **(description)**
What colour are blood platelets? **(description)**
How big are blood platelets? **(description)**
What shape do blood platelets have? **(description)**
How are blood platelets similar to white blood cells? **(similarity)**
How are blood platelets different from white blood cells? **(differences/contrasts)**
What are blood platelets used for? **(function)**

A. The blood

The **blood** is a fluid for transporting materials. It is a suspension of cells in a pale yellow liquid called plasma. Blood can be separated into ***blood cells*** and ***plasma***. The plasma makes up about 55% of the blood. It is mostly water, but also contains dissolved nutrients, wastes and other substances. The blood cells make up about 45% of the blood, and they include *red blood cells*, *white blood cells* and *blood platelets*.

Red blood cells contain red pigment and haemoglobin and are used for carrying oxygen. They are red, around 7-8 μm (1 μm = 0.001 mm) in diameter and have a biconcave disc shape. **White blood cells** protect us against diseases and kill bacteria that get into our body. They are colourless, irregular in shape and have a diameter of around 6-30 μm. Similarly, **blood platelets** are also colourless and irregular in shape, but they are smaller in size with a diameter of only around 1-4 μm. Blood platelets protect us by forming blood clots to stop blood loss from the damaged blood vessels.

Figure 8. A section of the reading handouts for revision of the circulatory system unit.

In the reading handouts adapted from the textbook, there were texts which were more coherent with clear genre structure, academic sentence patterns and academic vocabulary such as subject-specific words, general academic words and logical connectors which were highlighted with bold, italics and underlines to illustrate the different language features. In the handout for teachers, there were guiding questions designed according to the sentence patterns of different academic functions in the paragraphs so that the teacher may guide students to focus on functions such as defining, describing, classifying, exemplifying, etc. These various sentence patterns of academic functions and their corresponding examples in the text were further presented in PowerPoint slides (Figure 9) which Ms Wan explained to the students to raise their academic language awareness.

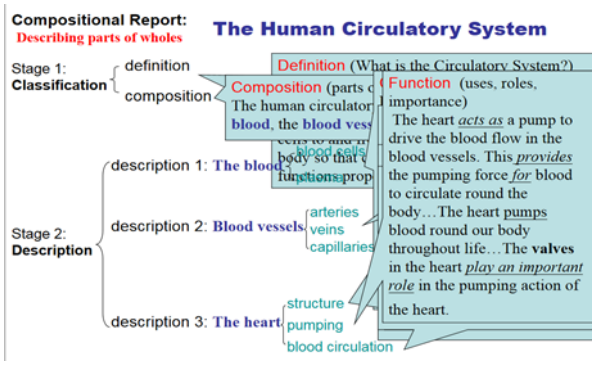


Figure 9. PowerPoint slides about the academic functions and examples in the circulatory system unit.

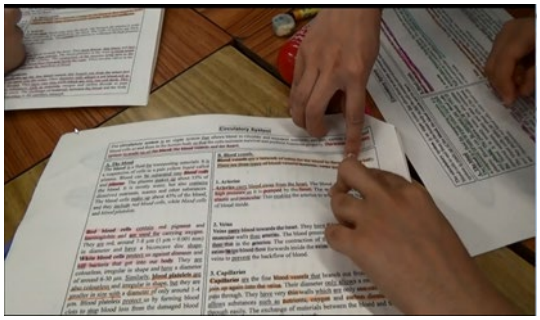


Figure 10. Collaborative learning based on the notes in the reading handout

During the revision lesson, Ms Wan drew students' attention to the different academic functions and their examples in the adapted text handouts. She then used the first paragraph of the handout; that is, the introduction about the circulatory system, as an example to demonstrate how to ask questions according to the academic functions of the sentence patterns in the paragraph. After students were able to identify the different academic functions and their corresponding sentence patterns, Ms Wan grouped the students and encouraged group members to ask each other questions according to the academic functions they could identify in different paragraphs in the reading handouts. Grounded in the handout, students read carefully, using markers to highlight the different academic functions and the relevant sentence patterns, and then started to ask each other questions in groups (Figure 10). Since students' English language proficiency and academic literacy varied largely among the class, those who were able to understand the text better helped their peers to unpack the meaning of the texts. They then repacked the information to answer the questions with the help of the sentence-making tables provided in the handouts. During the collaborative reading activity, Ms Wan allowed the more capable peers to explain the text to their less capable peers in

their first language, and then provided feedback and language prompts to help the students to express the sentences in English. According to the students in the interview, they found the reading handouts useful and used them for self-directed learning after class. In the last lesson, the students had a quiz about the unit which consisted of a puzzle word game about the key concepts and four short-answer questions about the circulatory system. The results of the papers indicated that students were able to write longer sentences and more logical answers to the open-ended questions compared with their previous quizzes.

Discussion

In this section, we shall discuss the results of the current study by focusing on meaning making by translanguaging and multimodal activities in the multilingual science classroom, the spaces of translanguaging and trans-semiotizing as well as the spaces for academic language development in the MEC-based curriculum genre.

Meaning making through translanguaging and multimodal activities in CLIL classroom

As Figure 2 shows, during the discussion of the worksheet exercises in the first lesson, meaning making was realised via multimodal activities (Danielsson, 2016; Lemke, 1998). For example, the introduction of the circulatory system, which was a labeling diagram exercise on the worksheet, was carried out through teacher-student interactions during which metaphoric language was used to describe similar structural characteristics between a city transportation system and a human circulatory system. The components of blood and their corresponding features, which were represented by a graphic organizer indicating classification of blood components and a table comparing the colour, shape and function of different blood cells, were demonstrated through a video accompanied by the teacher's verbal elaboration and sketches on the blackboard. The identification

of different blood cells displayed on a microphotograph was clarified by the teacher's drawing on the worksheet, gestures, together with verbal comparison of sketches on the blackboard. The different modes interplayed to facilitate meaning making during knowledge construction. As Danielsson (2016) pointed out, each mode (e.g., word, image and action) representing a concept has its special "meaning potential"; that is, the "affordance" that may convey a certain aspect of the conceptual meaning. For instance, visual images such as the diagram of the circulatory system provided information about the positions of the heart and vessels as well as their intertwining relations; while the arrows below the biconcave shape in the sketch on the blackboard also indicated the directions of light in a microscope. Gestures such as Ms Wan's finger circling in the air to show "circulate" and "going round and round", and her gesture of swirling the pen rapidly to illustrate the spinning of a centrifuge both added dynamic features to the meaning of the concepts. On the other hand, the verbal explanation and exemplification by the teacher's spoken language and written notes in the worksheet helped to clarify the abstract meaning or even correct misleading concepts (e.g., 'White blood cell is not white but colourless') which are unobservable and cannot be shown by either the visual or gestural modes (Kress, et al. 2001). The shifting between multiple modes during knowledge building did not just repeat the same meaning of the concepts, but multiply and amplify their meaning (Lemke, 1998) by highlighting the different characteristics and complexities in the concepts. The orchestration of different modes to elaborate on the same concept also helped to make meaning more dynamic and impressive. For example, during the discussion of the shape of red blood cells, the teacher's verbal description served as the background when she drew a biconcave shape on the worksheet projected on the screen, which became the foreground; that is, something attracting the audience's attention. However, when she gestured the "going down" action with her hands and simultaneously uttered the Cantonese words "凹咗", her drawing on the worksheet was backgrounded while the gesture and the verbal explanation became foregrounded (Kress, et al. 2001).

Such dynamic foregrounding and backgrounding interplayed to facilitate comprehension and consolidation of the meaning of the abstract concepts during the science lesson.

Observation of the lesson shows that, the presentation and discussion of concepts at different stages of the lesson was guided by designed scaffolding--the multimodal worksheet, and then implemented through spontaneous scaffolding (Gibbons, 2009; Lin, 2016). These include the interactions between the teacher and students in *flows* of translanguaging (Lemke, 2016; 2018; Lin, Wu & Lemke, 2019) and trans-semiotizing (Lin, 2015a; 2019), interweaving sounds (spoken English, Urdu and Cantonese), acts (speaking, drawing and writing on worksheets and blackboard), artefacts (worksheet, textbook and realia), meaning (mutual communication of concepts) and feeling (e.g., imagination of dizziness in the loss of blood) into *spaces* (Lin, Wu & Lemke, 2019) of information that connected events of now (blood and blood cells being discussed) and then (a student's blood donation experience and the teacher's experiment about blood in her previous practicum), here (discussion about blood cells in the classroom) and there (blood transfusion at Red Cross), concrete (city transportation system) and abstract (human circulatory system), colloquial (donut) and academic (biconcave) through verbal (i.e., spoken languages such as English, Urdu and Cantonese; and written texts such as the textbook and the student's English and Chinese notes on the worksheets), visual (i.e., videos, diagrams, graphic organizers, tables, photos, drawings and sketches, etc.) and actional (hand gestures, body movements and facial expressions) modes which interplayed seamlessly to expand the socially shared communicative repertoire (Lin, 2012; He, Lai & Lin, 2017), hence effectively facilitating meaning co-making in ongoing inquiry of the science knowledge.

Creation of spaces of translanguaging and trans-semiotizing and spaces for academic language/register use in MEC

An overview of the six lessons in the circulatory system unit (Table 1) shows not only the

various spaces of translanguaging and trans-semiotizing but also spaces for academic language development. While there seemed to be no fixed patterns within each specific space with seamless shifting between different modes and dynamic *flows* of translanguaging and trans-semiotizing across timescales and socio-ecological systems, the different lessons seemed to follow the stages of the MEC (Lin, 2015b; 2016) as a curriculum genre; namely, using multimodality to engage students in comprehension and discussion of the concepts about the content subject so that they were able to build up the thematic patterns (Lemke, 1990) of the unit; that is, the interconnected structures of semantic relationships between different thematic items, or concepts in the subjects. Then the teacher guided students to deconstruct texts and take notes to raise their awareness of academic language features by guiding them to identify the language features such as text-type structure, sentence patterns and functions, and academic vocabulary. For example, the first four lessons in the unit mainly focused on the presentation and practice of the basic content knowledge; for each lesson, students had rich experiential opportunities (e.g., video, images, experiments and games, etc.) to discuss and explore the relevant knowledge on the topic; and then had hands-on workbook exercises which involved some reading (e.g., decoding questions and identifying question prompts) and writing (e.g., answering questions in complete and grammatical sentences) exercises which help students to develop their academic literacy. By doing so, the teacher followed the MEC in each of her lesson so that the cycle of three stages reiterated in different lessons. Then after the key conceptual knowledge of the whole unit had been taught, in the fifth lesson Ms Wan guided students to have a general revision. She provided the students with adapted texts to demonstrate the useful language features so that they had an opportunity to reflect on the integration of academic language and content learning through guided reading and collaborative learning with peers. In the last lesson, the teacher assigned a quiz for students which consisted of a few open ended short-answer questions. This is a stage-three practice according to MEC which guides the students to entextualize their leaning experiences into

formal academic texts. These reiterating cycles of the MEC reflected Lemke's (1990) proposals for content-based instruction: "...Reasoning is combing the use of thematic pattern with the use of a rhetorical or genre structure pattern. One supplies the content, the other supplies the form of organization of the argument" (p. 122-123) and "Students will begin to grasp semantic and conceptual relations in colloquial language first. Then they will substitute scientific, technical terms for colloquial words" (p. 173). Namely, in multilingual classroom, teachers can allow students to make good use of their prior experiences, home culture and familiar languages to facilitate comprehension of the content subject (Cummins, 2015). Apart from content knowledge development, the teacher may also help students to foster the development of academic language awareness from first using colloquial everyday registers (e.g., "like a donut") to gradually shifting to technical and professional academic discourse (e.g., "a biconcave shape").

Table 1

Translanguaging and Multimodal activities in Multimodalities-Entextualization Cycle (MEC)

Lessons	**Multimodal Activities**	**Modes & Multimedia**	**TL-TS**
Lesson 1 Blood and Blood vessels	-T-Ss interaction -video watching -worksheet blank filling -workbook exercise	-textbook, worksheet, workbook, -video, diagram, table, photo, blackboard sketches, -gestures	**TL:** English; Urdu; Cantonese; Traditional written Chinese **TS**
Lesson2 Revision & Experiment	-Introduction of experiment apparatus and procedure -Experiment 1: observing blood vessels -worksheet exercise	-textbook, worksheet, workbook, -experiment materials and apparatus, -blackboard sketches, -gestures	**TL:** English and Urdu **TS**
Lesson3 Heart	-T-Ss interaction -observing/feeling organ models -worksheet blank filling -using teaching realia to imitate the blood flow direction and movement -workbook exercise	-textbook, worksheet, workbook, -diagrams, pictures -teaching realia, model of organs in human body	**TL:** English and Urdu **TS**

Lesson4 Revision & Experiment	-revision on blood circulation -Experiment 2: 1.dissecting an ox heart 2.measuring heartbeat, pulse and blood pressure -workbook exercise	-textbook, workbook, worksheet, -diagrams, pictures -experiment materials and apparatus, -stethoscope, sphygmomanometer, -gestures	**TL:** English and Urdu **TS**
Lesson5 Revision	-C+L games -complete the lyrics of a song "circulatory system" -guided reading -peer questioning -identifying and highlighting key language points -collaborative learning	-textbook, worksheet, text handouts, -diagrams, tables, graphic organizers -word games, song in video, PPT	**TL:** English and Urdu **TS**
Lesson 6 Quiz	-complete quiz questions -students receive scaffolding by teacher	-quiz paper -puzzle game	**TL:** English, Urdu and Cantonese **TS**

Note: TL: translanguaging; TS: trans-semiotizing

Conclusion

In this study, we analysed Ms Wan's CLIL lessons which provided practical examples of multilingual and multicultural CLIL classroom practices. From the students' active engagement; for example, their meaningful questions and co-construction of knowledge with the teacher as well as their worksheet notes and improved assignments, it can be seen that the science lessons about the circulatory system were effective. This did not seem to have been affected by the linguistic, cognitive or cultural diversity, as the teacher and students were not fettered by the "English only" rule of many traditional EMI classrooms; rather, the students were allowed to flexibly translanguage between English as the target language, Urdu as their first language and Cantonese as the most widely used language in the city. The different semiotic resources also allowed them to make meaning through translanguaging and trans-semiotizing in *flows* which started from the teacher to some students who co-constructed their information both in their mind (interpreting conceptual meaning), languages (spoken and written English, spoken Urdu, Cantonese and traditional written Chinese), gestures (trying to express ideas by gestures when lacking the vocabulary to express themselves) and physical

artefacts (their worksheet notes showing traces of the teacher's verbal explanations, sketches, drawing and words on the blackboard or the screen). And then through group work, these different sources of information were entrained further into the *flow* of meaning co-making among the peers in the learning community of the CLIL classroom (Lemke, 2016; Lin, Wu & Lemke, 2019). What was discussed in the science lesson may not only transcend the boundary of languages but also connect different time-scales (e.g., a student's blood donation before), physical environment (e.g., blood transfusion in Red Cross), or textbook content (e.g., the relationship between blood and DNA which will not be taught until Secondary 5). In this sense, meaning making in Ms Wan's multilingual CLIL classroom broke the limits and transcended the boundaries (Li & Lin, 2019) of traditional monolingual classrooms. The translanguaging and multimodal practices based on the MEC curriculum genre provided an example of lesson design to tackle the pedagogical challenges--planning and delivering content-and-language-integrated lessons, which confront CLIL teachers, especially pre-service teachers (He & Lin, 2018; Lo, 2020). Seeing that the diverse cultural and linguistic resources of Ms Wan's CLIL classroom had actually enriched the communicative repertoire shared by both the teacher and the South Asian minoritized students, and the learner diversity in the school may turn out to be beneficial rather than hindering students' content and language development, we propose that future CLIL programmes should change from a "Deficit-based Model" of teaching to an "Asset-Based Model" of teaching, mobilizing the multilingual and multimodal resources in the classroom with the MEC as a heuristic tool for planning CLIL lessons (Lin, 2020).

Notes

1.QEF Project. This paper is based on data from a Quality Education Fund (QEF) project funded by the Hong Kong Education Bureau (Project #EDB/QEF/2012/0483) awarded to Angel M. Y. Lin.

2. All names of teacher and students are pseudonyms in this article.

References

Baker, C. (2011). *Foundations of bilingual education and bilingualism* (5th ed.). Bristol: Multilingual Matters.

Cheng, M.M.W., & Gilbert, J.K. (2015). Students' visualization of diagrams representing the human circulatory system: The use of spatial isomorphism and representational conventions. *International Journal of Science Education*, *37*(1), 136–161. https://doi.org/10.1080/09500693.2014.969359

Coyle, D., Hood, P., & Marsh, D. (2010). *CLIL: Content and language integrated learning.* Cambridge: Cambridge University Press.

Cummins, J. (2015). Evidence-based TESOL: Teaching through a multilingual lens. Keynote speech presented at the TESOL Conference, Toronto, Canada, March.

Dalton-Puffer, C. (2018). Postscriptum: Research pathways in CLIL/Immersion instructional practices and teacher development. *International Journal of Bilingual Education and Bilingualism, 21(3),* 384‑387. https://doi.org/10.1080/13670050.2017.1384448

Danielsson, K. (2016). "Modes and meaning in the classroom–the role of different semiotic resources to convey meaning in science classrooms." *Linguistics and Education*, *35,* 88-99.

García, O. (2009). *Bilingual education in the 21st Century: A global perspective*. Oxford: Wiley.

García, O., & Li, W. (2014). *Translanguaging: Language, bilingualism and education*. Hampshire, U.K.: Palgrave.

García, O., & Lin, A. M. Y. (2017). Translanguaging in Bilingual Education. In O. García, A.M.Y. Lin, & S. May (eds.), *Bilingual and Multilingual Education, Encyclopedia of Language and Education* (pp.117-130). Dordrecht: Springer.

Gibbons, P. (2009). *English learners, academic literacy, and thinking: Learning in the challenge zone*. Portsmouth, NH:Heinemann.

Halliday, M. (2013). Languages, and language, in today's changing world. Seminar presented at the University of Hong Kong, Hong Kong, October

He, P., Lai, H., & Lin, A. M. Y. (2017). Translanguaging in a multimodal mathematics presentation. In Catherine M. Mazak, and Kevin S. Carroll (ed), *Translanguaging in higher education: Beyond monolingual ideologies.* 91-120. Bristol: Multilingual Matters.

He, P., & Lin, A.M.Y. (2018). Becoming a 'language-aware' content teacher: Content and language integrated learning (CLIL) teacher professional development as a collaborative, dynamic, dialogic process. *Journal of Immersion and Content-Based Education, 6*(2), 162–188.

He, P. & Lin, A.M.Y. (2019). Co-developing Science Literacy and Foreign Language Literacy through "Concept + Language Mapping". *Journal of Immersion and Content-Based Language Education, 7(2),* 261-288.

Karlsson, A., Nygård Larsson, P., & Jakobsson, A. (2019). Multilingual students' use of translanguaging in science classrooms, *International Journal of Science Education*, 41:15, 2049-2069, DOI: 10.1080/09500693.2018.1477261

Kress, G., Jewitt, C. ,Ogborn, J., & Tsatsarelis, C. (2001). *Multimodal teaching and Learning: The rhetorics of the science classroom*. London: Continuum.

Lemke, J. (1990). *Talking science: Language, learning and values*. Norwood, NJ: Ablex, .

Lemke, J. (1998). Multiplying meaning: Visual and verbal semiotics in scientific text. In J. R. Martin & R. Veel (Ed.), *Reading Science: Critical and Functional Perspectives on Discourses of*

Science, 87-113.
https://books.google.com.hk/books?id=rJuIfnu22RkC&printsec=frontcover&hl=zh-TW&source=gbs_ge_summary_r&cad=0#v=onepage&q&f=false

Lemke, J. (2016). Translanguaging and Flows. Unpublished research manuscript.

Lemke, J. (2018). Translanguaging and Flows. Webinar hosted at Simon Fraser University, Canada, 18 September 2018.

Li, W. (2018). Translanguaging as a Practical Theory of Language. *Applied Linguistics* 39 (1): 9‑30. doi:10.1093/applin/amx039.

Li, W. & Ho, W. Y. (2018). Language learning sans frontiers: A translanguaging view. *Annual Review of Applied Linguistics. 38.* 33-59. http://dx.doi.org/10.1017/S0267190518000053

Li, W. & Lin, A.M.Y. (2019). Translanguaging classroom discourse: pushing limits, breaking boundaries. *Classroom Discourse.10(5), 209-215.* https://doi.org/10.1080/19463014.2019.1635032

Lin, A. M. Y. (2012). Multilingual and multimodal resources in L2 English content classrooms. In Leung, C., and B. Street (ed). *"English"–A changing medium for Education* , 79-103. Bristol: Multilingual Matters.

Lin, A. M. Y. (2015a). "Egalitarian bi/multilingualism and trans-semiotizing in a global world." In In Wayne E. Wright, Sovicheth Boun, and Ofelia García (Eds.), *The handbook of bilingual and multilingual education 19-37*. West Sussex, UK: Wiley Blackwell.

Lin, A. M. Y. (2015b). Conceptualizing the potential role of L1 in content and language integrated learning (CLIL). *Language, Culture and Curriculum, 28* (1), 74-89.

Lin, A. M. Y. (2016). *Language across the curriculum & CLIL in English as an Additional Language (EAL) Contexts: Theory and practice*. Dordrecht: Springer. https://doi.org/10.1007/978-981-10-1802-2

Lin, A. M. Y. (2019). Theories of trans/languaging and trans-semiotizing: implications for content-based education classrooms. *International Journal of Bilingual Education and Bilingualism*, 1-12.

Lin, A. M. Y. (2020). From deficit-based teaching to asset-based teaching in higher education in BANA countries: cutting through "either-or" binaries with a heteroglossic plurilingual lens, *Language, Culture and Curriculum*, DOI: 10.1080/07908318.2020.1723927

Lin, A. M. Y. & He, P. (2017). Translanguaging as dynamic activity flows in CLIL classrooms, *Journal of Language, Identity & Education*, *16 (4),* 228-244.

Lin, A. M. Y. & Wu, Y. (2015). "May I Speak Cantonese?"‐Co-constructing a Scientific Proof in an EFL Junior Secondary Science Classroom. *International Journal of Bilingual Education and Bilingualism 18 (3),* 289-305.

Lin, A. M. Y. & Lo, Y. Y. (2018). The spread of English Medium Instruction programmes: Educational and research implications. In R.Barnard & Z. Hasim (Eds.), *English medium instruction programmes: Perspectives from South East Asian universities* (pp.104-115). London, England: Routledge.

Lin, A. M. Y., Wu, Y., & Lemke, J. L. (2019). "It Takes a Village to Research a Village": Conversations with Jay Lemke on Contemporary Issues in Translanguaging. In: S. Lau and S. Van Viegen Stille (Eds.), *Critical plurilingual pedagogies: Struggling toward equity rather than equality*. Switzerland: Springer.

Lo, Y.Y. (2020). *Professional Development of CLIL Teachers*. Singapore: Springer

Miles, M.B., & Huberman, A.M. (1994). *Qualitative data analysis: An expanded sourcebook.* Thousand Oaks, CA: Sage.

Rose, D., & Martin, J. R. (2012). *Learning to write, reading to learn: Genre, knowledge and pedagogy in the Sydney School*. Sheffield: Equinox.

Rothery, J. (1996). Making changes: Developing an educational linguistics. In R. Hasan, & G. Williams (Eds.), *Literacy in society* (pp. 86‐123). London: Longman.

Ruiz de Zarobe, Y. (2016). Introduction‐CLIL implementation: From policy-makers to individual initiatives. In Y. Ruiz de Zarobe (Ed.), *Content and language integrated learning: Language policy and pedagogical practice*. London: Routledge.

Swinney, R., & Velasco, P. (2011). *Connecting content and academic language for English learners and struggling students grades 2-6*. (pp. 27-44, 63). London, England: SAGE Ltd.

Van Leeuwen, T. (2005). *Introducing social semiotics*. New York: Routledge.

Selected Papers

Research Articles

Exploring Learner Beliefs in Self-Regulated Learning: A Case Investigation of an English Self-Study

Akiko Fukuda

Rikkyo University

Abstract

This study investigated how learner beliefs relating to self-regulated learning were characterized and whether these beliefs changed over time during English self-study. Self-regulated learning functions effectively when supported by motivational beliefs, and many studies have confirmed its role in classroom settings; however, self-study which is detached from teachers and grades has not been examined. Two university students (Kei and Takushi) participated in the research and were willing to learn extra English through self-study. Using online material, they began to learn English independently. Over seven months, they set and reflected on learning goals. Thirteen semi-structured interviews were conducted, and excerpts regarding the participants' beliefs were extracted and categorized using qualitative thematic analysis. Five learner belief elements seemed to explain attitudes towards self-regulated learning during self-study. Kei displayed strong learner beliefs regarding the value of learning, the self, and learning management, which appeared to be the core ideas underpinning his self-regulated learning. Takushi also had a strong belief in language learning and its value, but sometimes these interfaced negatively, leading him to experience difficulty in conducting self-study based only on his own motivation.

Keywords: self-regulated learning, English self-study, learner beliefs, case study

Introduction

In an English as a foreign language (EFL) environment such as Japan, studies have acknowledged that the amount of time spent learning English in school is inadequate for reaching ultimate second language (L2) proficiency (Hiromori, 2015). Most EFL learners who wish to improve their English language skills tend to study independently, outside the classroom. In such situations, they are required to study the L2 in a self-regulatory manner. Increasing research into the effect of self-regulated learning (SRL) on L2 learning has recognized that SRL predicts language proficiency (Meece, Anderman, & Anderman, 2006; Seker, 2016; Teng & Huang, 2019) and that SRL instruction has a significant influence on language skills (Lam, 2014; Little, 2009); however, few studies have addressed the learning environment outside the classroom. Besides, SRL has sub-components that move the learning cycle forward, including learner's fundamental motivational beliefs (Zimmerman, 2000). Learner beliefs can be a strong motivator in self-study where learners need to learn English voluntarily, but previous studies have exclusively examined self-efficacy beliefs in the process of SRL, not focusing on other types of learner beliefs. If a focus on SRL skills for self-study can meet the demands of L2 acquisition, a comprehensive understanding of learner beliefs can play its role in proceeding SRL. To investigate this topic, the present study explored learners' beliefs about SRL and reported on two Japanese EFL learners' cases, tracing their English self-study attitudes outside the classroom over seven months.

Literature Review

The Cycle of Self-Regulated Learning

The concept of SRL has become embedded in the field of learner psychology relating to second language acquisition (SLA) because SRL is considered to be a comprehensive model focusing on the

interrelationship between motivation and learning strategy (Dörnyei, 2005; Dörnyei & Ryan, 2015). SRL is defined as "the ways that learners systematically activate and sustain their cognitions, motivations, behaviors, and affects, toward the attainment of their goals" (Schunk & Greene, 2018, p. 1). Zimmerman (2000) proposed three phases of SRL: forethought, performance or volitional control, and self-reflection, which include sub-components, as shown in Figure 1. The forethought phase involves learners in setting goals and planning their learning strategically, including considering self-motivational beliefs, such as self-efficacy and outcome expectations. The performance or volitional control phase consists in controlling and managing learning with focused attention and self-observation. In the self-reflection phase, learners identify and evaluate the causes of good work (Zimmerman, 2000). It is generally understood that, the more self-regulatory learners become, the more they achieve in school (Pintrich & De Groot, 1990; Zimmerman, 1990). In an EFL context, researchers have acknowledged that SRL skills correlate with and influence L2 proficiency (Fukuda, 2018; Teng & Huang, 2019; Seker, 2016).

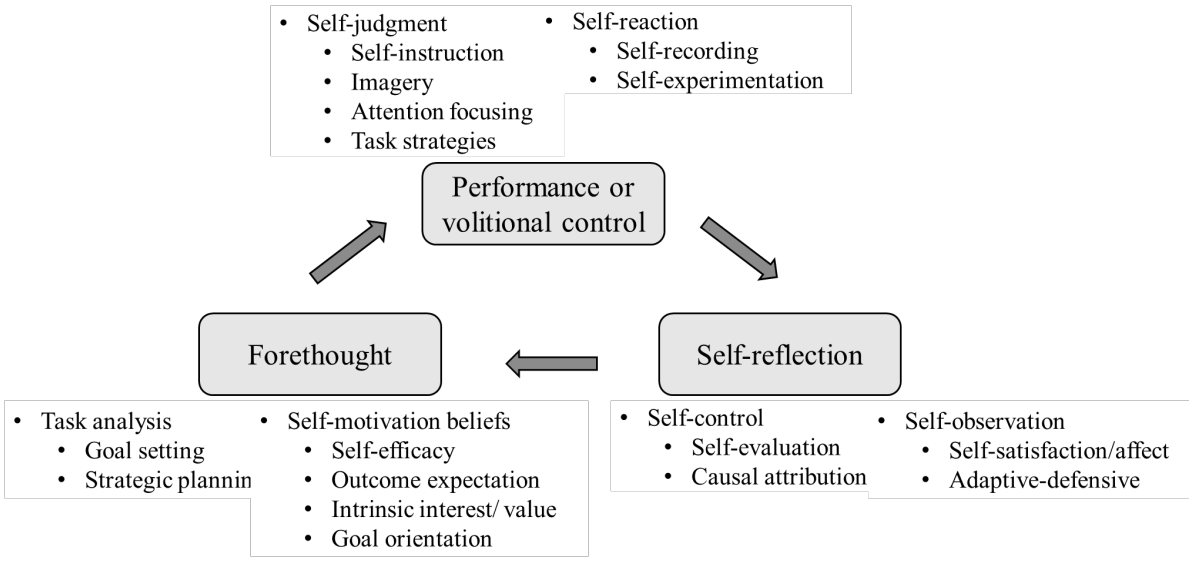


Figure 1. The cyclical model of self-regulated learning (based on Zimmerman, 2000).

Learner Beliefs

Learner beliefs were traditionally regarded as an accompanying element that supported rational argument, and were understood as stable innate individual traits (Dörnyei & Ryan, 2015); thus, many researchers attempted to classify different types of beliefs. One of the most widely used questionnaires, for example, was the Beliefs about Language Learning Inventory (BALLI) developed by Horwitz (1988), which assesses five elements of learner beliefs: the difficulty of language learning, L2 aptitude, the nature of language learning, learning strategies, and motivations and expectation. Early quantitative studies prioritized the investigation of the types of learner beliefs and their links with other factors using self-report questionnaires (e.g., Mori, 1999; Wenden, 1999).

Many studies have discussed learner beliefs, because of the ambiguity of their conceptualization; however, Cortazzi and Jin (1996) summarized the characteristics of learner beliefs as (1) referring to the nature of language and language learning and (2) emphasizing their social and cultural nature. Williams, Mercer, and Ryan (2015) widely generalized that beliefs represent "an acceptance or conviction that something is true" (p. 63). The current study regarded learner beliefs as referring, not only to the nature of language and language learning, but to any social and cultural factors underpinning the learning context—English language self-study.

Learner Beliefs in Self-Regulated Learning

Learner beliefs are especially concerned with the forethought phase of SRL, in that learners set goals, decide which learning strategy to use, and plan how to proceed based on their beliefs about L2 learning. Of the beliefs concerning the SRL process, self-efficacy is one of the most potent factors for enhancing the process (Mills, 2014). Cho and Kim (2019) confirmed that self-efficacy beliefs significantly related to self-regulatory skills for 173 Korean EFL learners, and the strength of self-efficacy beliefs and learners' SRL strategies differed significantly, depending on learners' English

proficiency. Even for English as a second language (ESL) children, the relationship between self-efficacy beliefs and SRL was found (Wang & Pape, 2005). Using a mixed-method case study involving questionnaires and interviews, Wang and Pape (2005) discovered that proficient English-speaking children tended to feel more efficacious, reporting wider and more numerous SRL strategies than less-proficient children.

Expanding self-efficacy beliefs, Zheng, Lian, Yang, and Tsai (2016) focused on conceptions (i.e., beliefs) about language learning and online self-regulation. By distributing two questionnaires about learner beliefs and SRL to 401 Chinese EFL learners, they demonstrated that two beliefs, understanding/seeing in a new way and memorizing, significantly predicted learners' self-regulated attitudes, while conceptions emphasizing success in examinations related negatively to online self-regulation. These quantitative studies provided us with clear-cut evidence of the relationship between learner beliefs and SRL, suggesting that learner beliefs about their confidence, outcome expectations (e.g., Scholer, Miele, Murayama, & Fujita, 2018; Zheng et al., 2016) and cognitive/metacognitive strategies (e.g., Cho & Kim, 2019; Law, Chan, & Sachs, 2008) have a positive relationship with SRL; however, such studies have hardly addressed the SRL cycle, despite the interaction between the three phases of SRL.

Qualitative Focus on Learner Beliefs and Self-Regulated Learning

The turning point for addressing learner beliefs came in the year 2000. Kalaja and Barcelos (2003) criticized the quantitative approach, which assumed that beliefs are stable and fixed mental representations, claiming that they are dynamic and situated (Barcelos, 2003, 2015). Ellis (2008) supported this argument, indicating that a qualitative method would be more suitable than a quantitative one for capturing trends in beliefs because beliefs can affect both the process and product of language learning. It is now generally understood that learner beliefs are complicated and formed

by interconnections between personalities, social contexts, and learning experiences.

Navarro and Thornton's case study (2011) revealed the interfaces between beliefs and self-directed learning[1] actions. Eighteen Japanese EFL learners participated in the course. The variety of instruments, such as self-reports, the guidance of self-directed language learning, advising sessions, and reflective journals, were combined. During the English course integrated with a self-directed language learning module, learner beliefs changed over time and that the interconnectedness of beliefs and self-directed actions played an important role in shaping further belief development. On the other hand, Kaypak & Ortactepe (2014), though not focusing on SRL, examined changes in beliefs about English learning through overseas exchange programs. They conducted mixed-methods research with surveys and journals, concluding that there was no change in beliefs before and after five months of studying abroad.

Whereas it is recognized that learner beliefs are formulated through the learning experience and social life (Aro, 2012), there seems no unified view, with some studies finding learner beliefs to be mutable and developing, and others finding no change. Further, although there have been studies focusing on learner beliefs related to SRL, less research has addressed them in the L2 self-study environment with a qualitative approach.

Self-Regulated Learning during Self-Study

Many studies have focused on the role of SRL in classroom activities; some of them have dealt with independent learning contexts, such as distance learning and computer-assisted language learning. For example, the relationship between self-instruction and SRL capacity in online courses (Bown, 2009), the effect of homework on achievements and performance in the classroom (Kominato, 2016), and the outcomes of study preparation for biology examinations (Sebesta & Speth, 2017) have been examined. Because of the overlapping roles of researchers and teachers/instructors,

most of the studies have recognized self-study conditions in classroom environments. Little research has explored English learning outside the classroom that is unrelated to class credits, assignments, teachers and instructors, and course grades; however, in circumstances where autonomous English learning, within and beyond the educational classroom, is encouraged in EFL countries, a focus on SRL skills to support self-study is vital for the effective acquisition of L2 (Hiromori, 2015). The current study defined self-study as a situation in which students learn English independently, regardless of their university courses, and focused on the relationship between SRL and learner beliefs in such a situation.

Problems and Research Questions

Learner beliefs, including self-efficacy, seem crucial for supporting the SRL cycle and its functioning, and SRL appears to be the core of L2 learning in self-study conditions. Even though learner beliefs and SRL are inseparably linked in various contexts, such as in classrooms, distance courses, skill-specific instruction, it has not elucidated how they relate in English self-study conditions. Additionally, more attention needs to be paid to investigating learner beliefs descriptively and dynamically because learner beliefs can be mutable and interrelated within the SRL cycle. The current study aimed to qualitatively explore the learner beliefs that Japanese EFL learners held about self-regulated language learning, how their beliefs related to SRL, and how these beliefs changed during English language self-study.

Method

Participants

The researcher visited compulsory English classes at a private university in Japan and recruited participants by announcing that the research sought students who wanted to improve their English

skills. The participants voluntarily joined the study, not knowing that the research dealt with the concept of SRL. This research was a part of a dissertation project: nine people participated in the research, and the two of them, both of whom were male university students, are the focus of the current paper in that they reported and showed contrasted learner beliefs throughout the study.

Kei[2] was 21 years old and majoring in business and marketing. He had been studying English for nine years but had never studied abroad. At the beginning of the research, he declared that his latest TOEIC score was 685. He joined the research in September 2018, and completed it in the middle of June 2019, after participating for eight and a half months.

Takushi was 18 years old and majoring in intercultural communication. He had been learning English for seven years, including overseas study in Australia for three weeks. He had Eiken Grade 2 at the beginning of the research, and his latest TOEIC score was 805. He began participating in September 2018 and finished in the middle of April 2019. His participation in the study was for about seven months.

Both participants had final examinations at their university at the end of January, and each of them traveled abroad during the fall holiday: Kei went to Bali and Takushi to the United States. After finishing the fall semester, Kei started full-time job hunting, which continued until early June; Takushi belonged to a volunteer club, so he undertook voluntary activities in Cambodia during the spring vacation.

Materials

The current research provided English language self-study conditions that had no relationship to university courses. To this end, online English material and learning logs were used.

Online material for learning English. The English online material, whose program was originally provided by the participants' university and free for the students and staff, was used in the

present study for English self-study. Considering the accessibility of the learning materials by all the participants, the establishment of suitable self-study conditions, and the convenience of the learning record, this online learning material was judged to be the most appropriate. It included four dimensions of English skills: vocabulary, grammar, reading, and listening, targeting TOEIC, TOEFL, IELTS, and general English learning. The participants could choose any content that interested them, accessing the material from any device, anywhere, and at any time.

Learning logs. Learning logs were distributed to the participants to record what they learned on certain days and the extent to which they were motivated and felt a sense of achievement. They wrote down how long they studied for and what they learned, then scored their levels of motivation and sense of accomplishment for that day. The learning log was mainly used to help them reflect upon their English learning, making the interviews more fruitful and facilitating the participants' retrospection.

Procedure and Analysis

Semi-structured interviews. Semi-structured interviews were adopted to focus on the SRL attitudes, reveal the interconnectedness of the learning environment with self-study, and enhance the learners' disclosure of their learning and self-refection. Although the current research expected the cyclical activities to follow the previously mentioned three phases of SRL theory, the interviews only provided an opportunity to investigate the two stages of forethought and self-reflection, due to the research design. Because the learners were supposed to learn English by themselves in the performance or volitional control phase, it was not monitored but occasionally reported in the interview retrospectively. A number of standardized questions were prepared, corresponding with the theoretical SRL framework (Zimmerman, 2000), as shown in Table 1, but the interviews were not limited to them.

Procedure. The interview sessions were conducted 13 times, biweekly. The interviews were arranged according to the participants' schedules and held in a private room at the university. Participants dropped in on their way to classes or clubs. To focus on the cyclical events of SRL, the interview sessions repeatedly required the learners to set goals and reflect on their SRL, to provide opportunities for forethought and self-reflection.

Table 1

Examples of the Prepared Questions for Each Phase

Forethought phase	Self-reflection phase
Q: Why did you set this goal?	Q: What were the good/bad points during this period?
Q: Do you expect that you will work well?	Q: Why did you (not) study well?
Q: How much are you motivated now?	Q: What was your motivation like?
	Q: To what extent did you feel a sense of accomplishment?

After receiving an explanation of the research objective and participation conditions, they described their English learning backgrounds and tried to set initial learning goals in the first interviews. During the intervals between the interviews, the participants studied English independently, using the online material, and completed their learning logs. After 11 SRL sessions, the learners had final interviews that enabled them to reflect on all the sessions (see Figure 2). Longitudinal research had the advantage of describing unique learners' characteristics as they relate to learner beliefs and SRL, and this approach allowed us to capture changes in beliefs during the research. This research attempted to establish a quasi-self-study condition and keep contact with learners to obtain data by repeating interviews and self-study period.

Theoretical concepts of beliefs and SRL had been concealed from the participants, whereas, information about them was implicitly elicited by the interviews. No instructions or advice were given to participants for setting goals or reflecting on their behaviors and motivations; therefore, the

ways in which they worked on their English learning varied. Even if they asked the researcher for advice or suggestions, they did not receive any; rather, they had to address the issues by themselves. The researcher restricted herself to listening attentively to the learners' talk, holding all the sessions as openly and freely as possible, and trying to pose the prepared questions gently in the course of natural conversation; thus, the participants could disclose anything about their daily lives, their moods, and their feelings and thoughts.

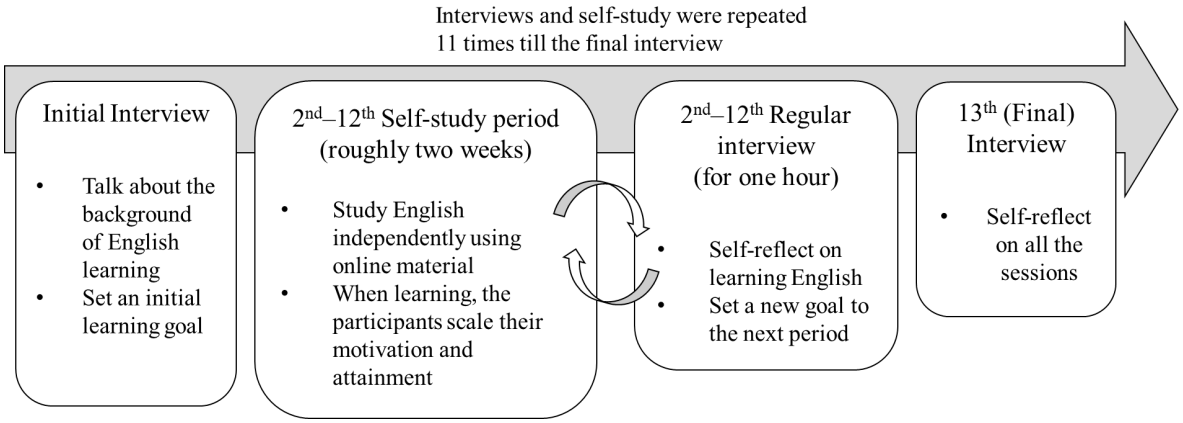


Figure 2. The research procedure.

Analysis. Thematic analysis (Tsuchiya, 2016) was applied to the interview data. All the interviews were transcribed and, thereafter, the researcher read the transcripts repeatedly. The transcribed data was then segmented, with one paragraph of conversation regarded as a single unit of coding. The units of coding were then categorized and labeled with thematic codes. The research aimed to focus, not only on beliefs about linguistic features, but any contextual beliefs; every mention of beliefs was extracted according to the definition of Williams, Mercer, and Ryan (2015). This work was done in a timely and repeated manner to ensure the reliability and validity of the data (Meriam, 1997).

The relationship between learner beliefs and SRL involved determining whether the beliefs

were reflected during the forethought and/or self-reflection phases of the study. Learner beliefs that were extracted by the thematic analysis, if different perspectives on the same beliefs were displayed in earlier or later interviews, were judged to have "changed."

Results and Discussion

The Relationship between Learner Beliefs and Self-Regulated Learning

The inductive thematic approach involved extracting five categories from the interview data: (1) motivational beliefs about the L2 (antecedents of participants' motivation to learn English before their self-study began); (2) learner beliefs about language learning; (3) learner beliefs about the value of examinations; (4) learner beliefs about the self; and (5) learner beliefs about learning management (participants' management of such aspects as time, the environment, and how to maintain their motivation to keep learning).

Kei's Case

Motivational beliefs about L2. Kei fundamentally thought that language learning would expand the world he could access. He also had a strong intrinsic interest in English, stating: "English is very cool, and the best way to improve my speaking skills is to communicate with native English speakers"[3] (1st interview). He greatly admired native speakers of English and cared about acquiring conversational skills. Another of his motives for studying English came from his interest in MBA programs abroad. After joining a company, he hoped to take advantage of the company's MBA program in the United States. He wanted to learn English to make his dreams of studying for an MBA come true.

Learner beliefs about language learning. Compared to Takushi, Kei thought more systematically about English learning, emphasizing two characteristic strategies for acquiring English

in light of his learning experience: continuity and skill integration. First, he understood that to acquire the L2, it was necessary to keep practicing regularly. Here are two extracts regarding his beliefs about effective ways of mastering English:

> (Extract 1) Let's say that you practice listening for two hours just one day during a week, or that you do it for 15 to 30 minutes every day for 2 hours per week in total. I think the latter is much better (Self-reflection, SR hereafter, 4th interview).
>
> (Extract 2) Vocabulary is limited, but listening and reading are infinite. I need to gain experience of reading and listening, so I can never have too much practice (SR, 8th interview).

As his remarks showed, Kei purposefully paid attention to learning English continuously; because he thought that the amount of practice would influence the proficiency the later achieved, he emphasized the frequency as well as the duration of practice. In addition, he held the belief that language skills cannot be separated, but must be acquired reciprocally; for example, he perceived that vocabulary was integral to writing and reading. He stated this with regard to ways of mastering vocabulary:

> (Extract 3) I don't enjoy studying using only vocabulary books: I hate them. I prefer using texts that contain integrated passages and vocabulary because I can learn the vocabulary through reading (1st interview).
>
> (Extract 4) Vocabulary can't be acquired without writing. I write messily on the back of used paper and, when it is full, I throw it away. Practice itself is important, so I don't care how neatly I write (SR, 3rd interview).

Vocabulary books are commonly used by students in Japan who are taking entrance examinations for universities. Most of them comprise English words and their equivalents in Japanese, with notes about pronunciation and intonation, and sometimes examples of word usages. Kei, however, declined to use such materials, preferring to choose his own material (mainly reading

texts with the necessary words integrated into them). It can be said that a strong belief appeared in his preference for these kinds of texts. Kei mainly practiced reading and listening sections in the research, using vocabulary and grammar sections supplementarily.

Learner beliefs about the value of examinations. In an EFL country such as Japan, the main objectives for studying English independently are connected to examinations. Kei expressed the value of English as it relates to assessment:

> (Extract 5) I trust the value of the nationwide mock exams more than the ranking of tests within schools. Anyone can get high enough scores on school tests if they prepare, so students with high GPAs [grade point averages] are not always smart. School tests are rather meaningless (forethought, F hereafter, 5th interview).
>
> (Extract 6) Taking TOEIC only in order to get a good grade is a waste of energy … English is meaningless unless you acquire it properly; thus, it shouldn't be achieved with a short-term purpose (SR, 4th interview) … I think that preparing for TOEIC to get good grades for job hunting doesn't make sense (F, 12th interview).
>
> (Extract 7) Taking TOEIC is for estimating how proficient I become. It is just one step toward ultimately talking confidently with foreigners and traveling abroad (Kei, 13th interview).

Kei regarded taking the examination as one way to confirm the development of his language skills. His long-term goal was to converse with foreigners and travel abroad fluently. Taking tests served to check how his English ability developed. Evidently, he did not see it as beneficial to take an examination simply to achieve a high score, as shown in Extract 7, because (in his words) learners tend to spend less time studying English when they set such short-term goals. He believed that English cannot be acquired easily, since he talked about the necessity of continuous learning, emphasizing that English takes a long time to acquire and good grades are a by-product of mastering English.

Learner beliefs about the self. Kei was particular about what he was to be. He mentioned that his motivation was not only to learn English but to frame his life policy. When he came to the interview session after a demotivated period, the researcher asked why he was able to learn English according to his predetermined goals, despite the lack of motivation, and he answered:

> (Extract 8) Those who are all talk and no action annoy me. I really hate liars and can't understand those who don't do what they say they will do. People who don't keep their word are the worst in the world. If you cannot do what you plan, it is only demotivation … so I keep my word. (SR, 7th interview) … Don't you think that people who don't keep their words for goals are uncool? I don't like all-talk-but-no-action people (SR, 10th interview).

This utterance encapsulated Kei's stand toward English self-study throughout the sessions. He was consistently true to his word, and this was evident in the forethought phase; that is, when he set his own goals to learn English through self-study, he tried to set achievable and realistic goals, because he did not want to lie to the researcher or to himself. If he failed to achieve his goals, he would have had to admit to being "all-talk-but-no-action," and he loathed such an attitude. It was clear that his beliefs spurred his learning into real action.

Learner beliefs about learning management. Goal-setting and controlling the learning circumstances were included in this category. Kei's policy for setting goals was to engage in English learning as often as possible:

> (Extract 9) When I'm free, I often drop in to the café and do something, such as preparing for job hunting or term papers. I buy a beverage, of course, so it would be a waste of money if I didn't do anything there. That's how I make myself work. This is my habit, and I integrate English study into this habit (SR, 9th interview).
>
> (Extract 10) I don't want to set goals that overwhelm me, but too easy goals are useless. I set a proper schedule that suits me (SR, 6th interview).

> (Extract 11) Goals need to be clear and easy to understand. Setting goals involves revising them periodically … Setting a vague goal such as "become better" makes no sense, because it is important to understand what I need to do. Once I confirm a small step, I can go on to the next goal (F, 10th interview).

Kei was very conscious of making English a habitual part of his life, because of his belief that English cannot be acquired in a short time (see Extracts 1 & 9). As one of his strategies for making himself study English, he mentioned going to a café. When he entered a café, he had to buy a drink to stay there; hence, he took advantage of this opportunity to do his schoolwork or prepare for job hunting. This habit made him finish his work, even when he was reluctant to do so; because if he did not do anything except fiddle with a smartphone, he would have wasted the money for a cup of coffee.

To keep learning English, he tried to be conscious of the content of his goals (see Extracts 10 and 11). There were three points, according to Kei, that made goals more effective: they should be specific, clear, and comprehensible. He often dialoged with himself, guessed that challenging goals might be the most appropriate for him, and then set goals with appropriate content, time, and measures. Because he believed that too easy or too difficult goals were meaningless, he was prudent in setting a new goal in each session.

Change in learner beliefs. In terms of the five categories of learner beliefs, Kei consistently adhered to what he talked about in the 13 interviews. After the final session, he successfully obtained a job with the possibility of studying for an MBA abroad. He remarked, "Three years before, I thought the guy who got a high TOEIC score was cool, but not anymore. Many of my friends failed in job hunting, even with high scores" (13th interview). Although his view of examinations might have changed throughout his university life, his beliefs did not change with regard to self-study.

Takushi's Case

Motivational beliefs about L2. Takushi thought of English as a tool for having conversations with foreign friends and teachers. He reported his desire to speak English fluently, but he was struggling with it, and he could not figure out the cause:

> (Extract 12) I don't know why, but I can get a high enough score on the listening test, so I'm worried about the future when I might not be able to do it suddenly. I can understand English, but I can't converse well … I have no idea why speaking and listening skills are so disconnected" (1st interview).

One of his interesting motives for trying self-study was because, in addition to acquiring English skills, he wanted to participate in the research. He was interested in pursuing a doctoral course, so he wanted to become familiar with research.

Learner beliefs about language learning. Takushi showed his unique opinion of how to acquire necessary words, stating that learning vocabulary is unnecessary. He said that if he came across unfamiliar words while learning, he would ignore them and focus on understanding the main ideas:

> (Extract 13) I think that remembering vocabulary is unnecessary for reading English passages because I don't get bad results when I don't study vocabulary … You can answer questions as long as you can understand sentences, despite ignoring unknown words (SR, 6th interview).

Nor did he think that practicing ways of answering and the correct spelling was necessary for him in language learning. For example, he mentioned that translation and revision technology has developed in recent years, implying that people no longer have to remember the spelling of words:

> (Extract 14) Practicing spelling is what high school students do, so I don't need it. Also, most writing must be done with laptops, so MS Word revises the spellings (F, 7th interview).
>
> (Extract 15) My vocabulary skill won't decrease, and my reading skill will be maintained, as

long as I practice listening. My grammar skill is … I think it will be okay with my spirit (F, 10th interview).

Based on his belief that he did not need to learn English systematically, he concluded that the necessary skills (i.e., reading, vocabulary, and grammar), except for listening, could be disregarded (Extract 15). Although he first learned vocabulary in this research due to its easiness, he eventually only completed listening tasks in the online material for the purpose of learning vocabulary.

Learner beliefs about the value of examinations. Takushi placed far more emphasis on taking examinations and achieving good results than Kei. The extracts below represent how he had studied for English examinations throughout his life. Characteristically, he understood examinations to be the equivalent of a *game* (Extract 16), with regard to the type of questions; for example, TOEIC uses multiple-choice answers for all 200 items. Takushi reported that he could choose options with a 25% probability of success, even if he answered unfamiliar questions; therefore, rather than studying to master English, he practiced narrowing down the four choices to just two, becoming more familiar with the format of the test than developing a specific skill (Extract 17).

(Extract 16) All the TOEIC questions have a choice of four answers, so it is easy to get points (1st interview) … I trained only on the strategic aspects of the TOEIC test. Anyway, my score will depend on whether my answers, based on the odds, are right or wrong (F, 12th interview).

(Extract 17) I feel that I can achieve an advantageous score for job hunting, even with overnight cramming. The job I'm interested in requires a TOEIC score of at least 860, so I will study for it someday, but I don't have to start it right now (SR, 4th interview).

(Extract 18) I want to use English well, because I want to be recognized as being intelligent. I don't know how to acquire English, but when I get a high score on TOEIC, I feel that I rank more highly as a human being (SR, 11th interview).

English examinations were likened to a game by Takushi, who believed that there was a knack

to passing English examinations. At the same time, he explained that his pleasure came, not from obtaining a high score and developing English proficiency, but from beating the people around him in terms of the examination scores (Extract 18); hence, his motivation declined during the self-study, which required beliefs about learning management.

Learner beliefs about the self. One of the most remarkable differences between the two learners was what they thought about themselves in relation to their learning of English. While Kei recounted firm beliefs about himself, Takushi did not mention himself at all. He described his environment, the influence of his friends, and the phenomenon of his language learning, but he did not clearly state his own beliefs.

Learner beliefs about learning management. Kei managed his environment to keep learning English, but Takushi's behavior depended on examinations. Takushi confided, in the final session, that he no longer wished to study English, as described in Extract 19. Given this change in his viewpoint, he could not imagine what he would be doing seven months later and could formulate suitable goals. He preferred extrinsic to intrinsic rewards, which gave him a much stronger sense of studying English to prepare for examinations; for example, he claimed that the researcher should have set conditions, with those who achieved the highest scores on the TOEIC exam at the end of the research obtaining a greater reward for beating others. This was his motivation for conducting English self-study: not to acquire English, but to beat his fellow students. For Takushi, English seemed to be a tool for gaining supremacy over others, as well as for communicating with them:

> (Extract 19) I had no idea of the necessary amount of English to study because I had no concrete goals … I couldn't imagine the future or set goals, even for seven months ahead. I couldn't estimate my progress, so I didn't try to set any objectives (13th interview).
>
> (Extract 20) If you had given me a goal such as "Improve your TOEIC score by 50 points", I could have decided on my goals and been motivated. To make efforts to accomplish what

> others dictate is good for me. Someday, when I become a member of society, I will have to work on this kind of self-study. So far, I have forced myself to study by applying for certification exams because it is an easier way for me to compete with others. (13th interview)

He used his train commute to study vocabulary and listening because he could use his smartphone to pass the time; however, this was not new learning for him, which made him more reluctant to practice English self-study.

Change in learner beliefs. Throughout the research, Takushi's beliefs about language learning strategies seemed to change. He gradually realized the importance of an awareness of grammar:

> (Extract 21) Learning English is about getting good grades in junior or high school, but since entering university, I have realized that English is a communication tool for conversations with foreign friends and teachers. I regret my study focus—concentrating on grammar … I'm especially bad at speaking and listening. My grammar knowledge is sufficient, but I think it is unnecessary to study abroad (1st interview).

Takushi originally wished to study abroad and wanted to take advantage of voluntary activities to visit foreign countries. Considering the environment around him, in which he was majoring in intercultural communication alongside many students from overseas, he noticed that English was a communication tool for conversing with these international students. At the same time, he realized and regretted that he devoted too much time and energy to acquiring English grammar, reporting that his priority, when learning English, was to develop listening and speaking skills for fluent communication.

However, after going abroad as a volunteer, his learner beliefs changed dramatically. He realized that grammar is the basis for constructing sentences and having conversations with his friends:

> (Extract 22) Most of my friends told me that my study style, focusing only on grammar and test

> scores, was a waste, and I thought that myself, but I noticed the usefulness of grammar a bit in Cambodia. I thought that grammatical skill wasn't needed in a conversation, but now I think it is important (SR, 12th interview).

Takushi had a conversation with some local friends in Cambodia, during which they told him that they spoke fluently, but that their grammar was weak; therefore, they respected his grammatical ability, since he could make a conversation fruitful. This experience led Takushi to notice the importance of a grammatical basis for communicating in English. The change came about, not because of the self-study, but due to his volunteer experience, which did not seem to relate to self-study.

Summary. Both learners joined the current study with reasonable motivation; however, they ran their self-study in a very different way with the same material and context. Kei tried self-study because he originally had a positive and intrinsic interest in using English and especially wanted to become a fluent English speaker. After starting self-study, he had cared about the number and quality of his goals, frequently evaluating whether he would achieve them or be overwhelmed. Obviously, his attitude to goal-setting and motivating himself was supported by his beliefs. His goals guided his mastering of English, which he believed to be a process of creating small steps that helped him to achieve his ends more easily. He mentioned that goals needed to be clear and easy to understand, reflected on how he could achieve his aims, and revised goals that were not difficult enough. This is the very process that many self-regulated learners have traced (Zimmerman, 2002). This belief in learning management seemed to be promoted by his belief in the value of examinations, which he perceived a means to assess his L2 development, and by his beliefs about himself. Kei believed strongly in his motto: "If you give your word, you must keep it," which enabled him to foresee what he could do and reflect on how he behaved. Eventually, these consistent and firm beliefs played a role in his SRL of English by self-study.

Takushi was first motivated to speak English because English appeared to be a necessary tool for communication with people around the world. Once he joined the study, he experienced difficulty in setting learning goals because his motivation for self-study depended on getting better scores than the people around him, and he could not achieve such an aim in an individual learning situation. He could have set examination-oriented goals to advance his self-study, but he did not do so. It caused him to feel that there was no benefit to taking tests in the period of the research. His beliefs in the value of learning English seemed to have a negative impact on his goal-setting. Additionally, he did not have strong beliefs about himself, in contrast to Kei. These elements might have affected the forethought phase of SRL: his self-reflection was not likely to lead to goal-setting in the forethought phase. He believed that goals should be set by others, such as the requirement to take a class or achieve a good grade, and such extrinsic goals motivated him. He thought himself capable of studying English without setting goals; hence, it can be concluded that he was not self- but other-regulated.

General Discussion

The present study aimed to explore the relationship between learner beliefs and self-regulated language learning by analyzing two case studies. The results showed contrasting learner beliefs regarding L2 self-study and illuminated five categories of beliefs: motivation, language learning, the value of examinations, self-beliefs, and learning management. Learner beliefs and SRL clearly related to each other, regardless of the characteristics involved, which was consistent with previous studies hypothesizing that learner beliefs are indispensable in launching the SRL cycle (Navarro & Thornton, 2011; Zheng et al., 2016). Holding strong beliefs about the self can play an important role in SRL, but only Kei displayed such beliefs in the research. Being conscious of the self is a core idea underpinning SRL because every phase of SRL requires learners to focus on themselves in order to

ascertain whether their goals are suitable, both personally and situationally, and whether their learning performance is effective for attaining their goals (Zimmerman, 2002). Kei's habit of continually evaluating himself and his learning process strongly related to the concept of the L2 Motivational Self System (Dörnyei, 2009). Kei considered both an "ideal" and "ought-to" L2 self, in that he wanted to speak English like a native English speaker (ideal L2 self) and, once he had decided this, he thought that he should achieve it (ought-to self). Yashima, Nishida, and Mizumoto (2017) confirmed that both kinds of L2 selves can predict intended effort, and ideal L2 selves are influenced by beliefs about the communication orientation in L2 learning. This study showed that beliefs about the self may work as precursors for setting goals and be linked with the whole SRL cycle, as well as with making an effort to develop L2 learning habits.

With respect to learner beliefs about the value of examinations, both the two learners gave weight to them; however, what was obviously different between them was that Kei regarded them as a barometer of his development of English proficiency, while Takushi regarded them as a game enabling him to dominate others. This seems partially correspond to the results of Zheng et al. (2016), which suggested that an examination-oriented belief had a negative correlation to SRL. In an environment where initiative plays an important role in self-study, such as in this particular research context, judging the value of learning is crucial rather than being examination-oriented or not. In other words, understanding "why I am doing it" can enhance SRL for self-study, which is independent of teachers, grades, and course assignments in school. Takushi did not find such meaning in self-study because his powerful motivation depended on competition with other learners.

In terms of changes in learner beliefs, the present study found that there were hardly any changes in beliefs before and after the self-study sessions. Only Takushi showed the transformation of his grammatical orientation; however, this seemed to be greatly influenced by his experience of going to Cambodia as a volunteer. This result seemed to support the idea that a person's beliefs

cannot be transformed easily, as Kaypak & Ortactepe's (2014) study illustrated with regard to unchanged beliefs even pre- and post-study abroad periods. Navarro & Thornton's (2011) study identified a significant development of learner beliefs, asserting that both beliefs and self-directed language learning can be mutable through interaction with instructors. Similarly, Barcelos (2015) pointed out that beliefs are not stable, but dynamic and influential, and are formed by learners' experiences. It might, therefore, be worth lengthening the research period to trace changes in learner beliefs about SRL.

Last, SRL could be based on individual learners' intentions. As previous studies have mentioned, the cycle and strategy of SRL are teachable, and learners can acquire SRL skills through explicit instruction about their mechanisms (Lam, 2014; Little, 2009). The current study did not provide any direction or advice about SRL to the learners; nevertheless, Kei was encouraged to learn English autonomously by monitoring his goals and level of attainment. Takushi paid little attention to himself in the first interview, being far more conscious of other factors, such as the online material and the personalities of his teachers and friends; however, he gradually began to consider why he did not have goals. This indicated that less self-regulated learners might benefit from an opportunity to look at themselves objectively through self-study. Another possibility is that the participants took the opportunity to learn SRL activity implicitly through the influence of biweekly interviews, during which they could engage in reflecting on their self-study and setting goals for the forthcoming session. The researcher always invited the learners to sit comfortably and chat, which should have created plenty of opportunities for them to reflect on themselves. These calm situations could have contributed to the participants reporting that they did not notice the SRL cycle or understand the concept until the end of the research. SRL may thus be acquirable by learners without them necessarily being taught it, providing that they are fully aware of their learning attitudes, set goals, and reflecting on them.

Conclusion

In conclusion, the current research confirmed that learner beliefs were certainly embedded in SRL during English self-study. Learner beliefs about the value of language learning, the self, and learning management were seen as parts of the SRL cycle, with learners foreseeing the self-study by setting goals and motivating themselves. Although most studies have concerned SRL in classrooms, this research used a self-study setting, which did not relate to teachers, academic classes, or grade point averages (GPAs). It could be claimed that only part of the pure SRL situation requiring learners to learn the L2 in a self-regulatory manner was observed; however, there are two other limitations in this study. First, the research could only speculate on the relationship between beliefs and SRL, and the process behind changes in learner beliefs was not identified. More longitudinal research (over a year or a much more extended period) should be conducted to focus on the accumulation of beliefs about language learning. Learners' backgrounds should be examined more closely to address the influence of beliefs on other dynamic and complicated processes of L2 learning. Second, self-selection bias may have existed in that the decision to participate in the research was left entirely up to the participants, representing that the participants might show only a unique case out of the population. However, because the study was designed to purse the attitude in long-term English learning, it was reasonable that learners who were willing to learn English joined the research. The present research emphasized examining learner beliefs, so the complete SRL cycle was not highlighted. By triangulating a quantitative and qualitative approach, future research could accurately describe the SRL cycle, clarifying transformations in perspectives about SRL from pre- to post-English self-study.

Acknowledgment

I appreciate the efforts of the three anonymous reviewers in providing considerate and encouraging comments and suggestions. I am also grateful to Dr. Satomi Takahashi, professor at Rikkyo University, for her thoughtful and insightful support for the research.

Notes

1. Self-directed learning is a similar concept to SRL, in that both include goal-directed and self-organized learning, but self-directed learning is broader than SRL and features the learning environment, whereas SRL features learner characteristics (Loyens, Magda, & Rikers, 2008).
2. Names have been changed to ensure anonymity.
3. The interviews were conducted in Japanese. The transcriptions were translated into English and proofread by a native English speaker.

References

Aro, M. (2012). Effects of authority: Voicescapes in children's beliefs about the learning of English. *International Journal of Applied Linguistics, 22*(3), 331–346. https://doi.org/10.1111/j.1473-4192.2012.00314.x

Barcelos, A. M. F. (2003). Researching beliefs about SLA: A critical review. In P. Kalaja, & A. M. F. Barcelos, (Eds.), *Beliefs about SLA: New research approaches* (pp. 7–33). Dordrecht: Springer Netherlands.

Barcelos, A. M. F. (2015). Unveiling the relationship between language learning beliefs, emotions, and identities. *Studies in Second Language Learning and Teaching*, *5*(2), 301–325. https://doi.org/10.14746/ssllt.2015.5.2.6.

Bown, J. (2009). Self-regulatory strategies and agency in self-instructed language learning: A situated view. *Modern Language Journal, 93*(4), 570–583. https://doi.org/10.1111/j.1540-

4781.2009.00965.x

Cho, Y. A., & Kim, Y. (2019). The relationship between self-efficacy beliefs and self-regulated learning strategies in Korean EFL learners. *The Linguistics Association of Korea Journal, 27*(3), 53–74. https://doi.org/10.24303/lakdoi.2019.27.3.53

Cortazzi, M., & Jin, L. (1996). Cultures of learning: Learning classrooms in China. In H. Coleman (Ed.), *Society and the language classroom* (pp. 169–203). Cambridge: Cambridge University Press.

Dörnyei, Z. (2005). *The psychology of the language learner: Individual differences in second language acquisition.* NY: Routledge.

Dörnyei, Z. (2009). The L2 motivational self system. In Z. Dörnyei & E. Ushioda (Eds.), *Motivation, language identity and the L2 self* (pp. 9–42). Clevedon: Multilingual Matters.

Dörnyei, Z., & Ryan, S. (2015). *The psychology of the language learner revisited.* NY: Routledge.

Ellis, R. (2008). Learner beliefs and language learning. *Asian EFL Journal, 10*(4), Conference Proceedings, 7–25.

Fukuda, A. (2018). The Japanese EFL learners' self-regulated language learning and proficiency, *Journal of Pan-Pacific Association of Applied Linguistics, 22*(1), 65–87. https://doi.org/10.25256/paal.22.1.4

Hiromori, Y. (2015). *Eigogakushu no mekanizumu: Dainigengoshutokukenkyu ni motodzuku kokatekina benkyoho* [The mechanism of English learning: An effective way to learn based on the second language acquisition research]. Tokyo: Taishukanshoten.

Horwitz, E. K. (1988). The beliefs about language learning of beginning university foreign language students. *The Modern Language Journal, 72*(3), 283–294. http://dx.doi.org/10.1111/j.1540-4781.1988.tb04190.x

Kalaja, P., & Barcelos, A. M. F. (2003). *Researching beliefs about SLA: New research approaches*.

Dordrecht: Springer Netherlands.

Kaypak, E., & Ortactepe, D. (2014). Language learner beliefs and study abroad: A study on English as a lingua franca (ELF). *System, 42*(1), 355–367. https://doi.org/10.1016/j.system.2014.01.005

Kominato, A. (2016). The impact of homework on self-regulation and English self-efficacy among Japanese underprepared college students. *KATE Journal, 30*, 43-56. https://doi.org/10.20806/katejournal.30.0_43

Lam, R. (2014). Understanding EFL students' development of self-regulated learning in a process-oriented writing course. *TESOL Journal, 6*(3), 527–553. https://doi.org/10.1002/tesj.179

Law, Y., Chan, C. K. K., & Sachs, J. (2008). Beliefs about learning, self-regulated strategies and text comprehension among Chinese children. *British Journal of Educational Psychology, 78*(1), 51–73. https://psycnet.apa.org/doi/10.1348/000709907X179812

Little, D. (2009). Language learner autonomy and the European language portfolio: Two L2 English examples. *Language Teaching, 42*(2), 222–233. https://doi.org/10.1017/S0261444808005636

Loyens, S. M. M., Magda, J., & Rikers, R. M. J. (2008). Self-directed learning in problem-based learning and its relationships with self-regulated learning. *Educational Psychology Review, 20*(4), 411–427. http://dx.doi.org/10.1007/s10648-008-9082-7

Meece, J. L., Anderman, E. M., & Anderman, L. H. (2006). Classroom goal structure, student motivation, and academic achievement. *Annual Review of Psychology, 57*, 487–503. https://psycnet.apa.org/doi/10.1146/annurev.psych.56.091103.070258

Meriam, S. B. (1997). *Qualitative research and case study in applications in education.* San Francisco: Jossey-Bass.

Mills, N. (2014). Self-efficacy in second language acquisition. In S. Mercer & M. Williams (Eds.), *Multiple perspectives on the self in SLA* (pp. 1–22). Bristol: Multilingual Matters.

Mori, Y. (1999). Epistemological beliefs and language learning beliefs: What do language learners

believe about their learning? *Language Learning, 49*(3), 377–415. https://doi.org/10.1111/0023-8333.00094

Navarro, D., & Thornton, K. (2011). Investigating the relationship between belief and action in self-directed language learning. *System, 39*(3), 290–301. http://dx.doi.org/10.1016/j.system.2011.07.002

Pintrich, P. R., & De Groot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology, 82*(1), 33–40. https://psycnet.apa.org/doi/10.1037/0022-0663.82.1.33

Schunk, D. H., & Greene, J. A. (2018). Historical, contemporary, and future perspectives on self-regulated learning and performance. In D. H. Schunk, & J. A. Greene (Eds.), *Handbook of self-regulation of learning and performance* (pp. 1–15). NY: Routledge.

Scholer, A. A., Miele, D. B., Murayama, K., & Fujita, K. (2018). New directions in self-regulation: The role of metamotivational beliefs. *Current Directions in Psychological Science, 27*(6), 437–442. https://psycnet.apa.org/doi/10.1177/0963721418790549

Seker, M. (2016). The use of self-regulation strategies by foreign language learners and its role in language achievement. *Language Teaching Research, 20*(5), 600–618. https://doi.org/10.1177%2F1362168815578550

Sebesta, A. J., & Speth, E. B. (2017). How should I study for the exam? Self-regulated learning strategies and achievement in introductory Biology. *CBE Life Science Education, 16*(2), ar30, 1–12. https://dx.doi.org/10.1187%2Fcbe.16-09-0269

Teng, F., & Huang, J. (2019). Predictive effects of writing strategies for self-regulated learning on secondary school learners' EFL writing proficiency. *TESOL Quarterly, 53*(1), 232–247. https://doi.org/10.1002/tesq.462

Tsuchiya, M. (2016). *Thematic Analysis.* Kyoto: Nakanishiya publisher.

Wang, C., & Pape, S. (2005). Self-efficacy beliefs and self-regulated strategies in learning English as a second language: Four case studies. *CATESOL Journal, 17*(1), 76–90.

Wenden, A. (1999). An introduction to metacognitive knowledge and beliefs in language learning: Beyond the basics. *System, 27*(4), 435–441. https://doi.org/10.1016/S0346-251X(99)00043-3

Williams, M., Mercer, S., & Ryan, S. (2015). *Exploring psychology in language learning and teaching.* Oxford: Oxford University Press.

Yashima, T., Nishida, R., & Mizumoto, A. (2017). Influence of learner beliefs and gender on the motivating power of L2 selves. *The Modern Language Journal, 101*(4), 691–711. https://doi.org/10.1111/modl.12430

Zheng, C., Liang, J., Yang, Y., & Tsai, C. (2016). The relationship between Chinese university students' conceptions of language learning and their online self-regulation. *System, 57*, 66–78. http://dx.doi.org/10.1016%2Fj.system.2016.01.005

Zimmerman, B. J. (1990). Self-regulated learning and academic achievement: An overview. *Educational Psychologist, 25*(1), 3–17. https://doi.org/10.1207/s15326985ep2501_2

Zimmerman, B. J. (2000). Attaining self-regulation: A social cognitive perspective. In M. Boekaerts., P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 13–39). San Diego: Academic Press.

Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory into Practice, 41*(2), 64–70. https://doi.org/10.1207/s15430421tip4102_2

Re-Examining CLT: What Does It Mean to Be Communicative?

Chris Hastings

Aichi Prefectural University

Stephen Clarke

Nagoya College

Abstract

This paper argues that the L2 language teaching field has struggled to implement communicative language teaching (CLT) as it was originally conceived because theorists and practitioners have not reflected on the concept of communication deeply enough. As a result, the term CLT has come to mean very different things to different people. In order to move the field forward, the authors recommend that practitioners adopt VanPatten's (2019) definition of communication. This paper first explains VanPatten's definition in detail and gives examples of what kinds of practices and materials are communicative and uncommunicative. Next, the paper reviews other scholars' criteria for judging communicativeness and discusses their relative merits in comparison to VanPatten's definition. Finally, the paper states why VanPatten's definition of communication, although it may appear narrow and constraining, is so necessary for current language teaching.

Keywords: communicative language teaching, communicative tasks

Introduction

Since its inception over forty years ago, the concept of communicative language teaching (CLT) has been surrounded by confusion. It has always been a vague term that has been applied to many

different practices. As far back as 1982, Ellis (1982) noted that "the term communicative has no clearly understood and received meaning when it is applied to language teaching" (p.73) and Harmer (1982) bemoaned the fact that "everything is 'communicative' these days" (p.164). Nearly thirty years later, Littlewood (2011, p.541, in Arnold, Dörnyei and Pugliese, 2015, p.5) claimed that a "recurrent comment about communicative language teaching is that nobody knows what it is." This lack of specificity applies to the Japanese context as much as anywhere else. When asked to define CLT, Japanese teachers of English named such features as "the need for communication, self-expression, exchanging opinions in English, understanding English utterances, not worrying about grammar, guessing from contexts and general comprehension" (Sakui, 2004, p.159). From such a list, it is difficult to see how these different points could come together into any kind of coherent definition. Since teachers do not seem to know precisely what CLT means, one must wonder why the term has continued to be so widespread. One reason for its omnipresence in the field is due to the influence of publishers, who advertise their textbooks as being 'communicative' in order to boost sales. It may also suit teachers to adopt such a vague term because it allows them to make only minor changes to traditional practice.

This paper attempts to answer the question: 'What does it mean to be communicative?' In particular, it focuses on how teachers can evaluate language teaching materials and the discourse of their classrooms to judge whether these are 'communicative' or not. The authors support the position advocated by the second language acquisition (SLA) scholar Bill VanPatten that the work of theorists and practitioners in our field who use the term 'communicative' needs to be informed by a working definition of communication (VanPatten, 2017). However, many teachers struggle when asked to give a definition of the term 'communication' (VanPatten, 2017). Indeed, the authors of this paper were equally unable to define what communication is clearly before being acquainted with VanPatten's work. Given that it is standard practice in any academic endeavour for researchers and

theorists to define their terms clearly, this admission ought to promote concern and reflection. The field of second language teaching, however, has progressed with little awareness of the importance of having a working definition of communication, a concept that is referred to so frequently in everyday life that its meaning is taken for granted. This paper contends that many researchers have not thought deeply enough about what communication actually entails when they have used the term 'communicative' in relation to teaching. If we define our terms carefully, it will be possible to define CLT much more narrowly, and we believe this will help to move the field forward.

One of the original proponents of CLT, Sandra Savignon, has argued that the "essence of CLT is the engagement of learners in communication to allow them to develop their communicative competence" (Savignon, 2002, p. 7). In other words, "communication should not just be the goal of CLT, it should be the process of instruction itself" (Thornbury, 1998). This view of communication as both a means and as a goal of instruction (Pica, 2000) has been criticised (e.g. Celce-Murcia, Dörnyei, and Thurrell, 1997, Arnold et al. 2015), but the present authors believe it to be incontrovertible. Learners learn to communicate by engaging in communicative acts. Therefore, teachers who aim to promote proficiency in their classrooms (as opposed to test preparation, for example), are compelled to provide opportunities for genuinely communicative interaction in their classroom. In order to plan for this, we need to know precisely what kind of materials and practices can promote authentic communication. Equally, we need to know which parts of current practice are not truly communicative and then reconsider their value.

In this paper, we will first present VanPatten's (2019) definition of communication and discuss how it can successfully be used to judge the communicativeness of teaching materials and classroom discourse. As far as we know, this definition has not been reviewed in the literature before, and neither have we found an alternative definition with which to compare it. However, there have been alternative proposals in the literature of ways to categorise teaching activities as communicative or

otherwise. We will present a review of some of these examples and compare them with the adoption of VanPatten's (2019) definition of communication. Although Ellis (1982) argued that the term 'communicative' could also be applied to syllabus design, we believe it will be sufficient to identify the core aspects of CLT at the level of classroom practices. Our position is that the communicativeness of L2 teaching can best be distinguished by analysing the activities or tasks that learners are required to perform in a particular class.

VanPatten's Definition of Communication

VanPatten (2019), proposes a definition that is taken, in part, from the work of Sandra Savignon (1997): "Communication involves the expression, interpretation and sometimes negotiation of meaning in a given context for a given non-linguistic purpose." This definition can be used to judge whether genuine communication takes place in our classrooms and also whether teaching materials would likely promote it. Let us consider some examples of L2 teaching materials and classroom discourse.

If, for example, a student responds to the teacher's question, "What is the weather like today?", would this be a 'communicative' exchange? The answer is almost certainly 'no', because this exchange does not meet the criteria for communication set out in the definition. In the posing and the answering of the question, the expression and interpretation of meaning do occur, but the purpose of the question is to elicit a particular set of linguistic items (a description of the weather). Since both the teachers and the learners already know what the weather is like, nobody is discovering new information. The goal of the exchange, therefore, is to practice language use, so we would describe it as a linguistic goal. However, genuine communication always has a non-linguistic goal. VanPatten (2017) asserts that these goals can either be psycho-social, which involve the development and maintaining of human relationships, or cognitive-informational, which involve the exchanging of

information. For reasons that will be explained later, cognitive-informational goals are suitable for language classrooms. A third goal, entertainment, could also be added, but this is best left to activities outside the L2 classroom.

VanPatten's (2019) definition can be employed to analyse the communicativeness of materials aimed at any proficiency level and involving any of the four skills. Let us imagine listening materials, which involve learners listening to audio recordings and answering questions on the content in a textbook. Could we describe these activities as 'communicative'? The answer is 'yes', if important conditions are met. When listening, students are interpreting meaning that is being expressed in the recording. If the teacher pauses the recording and a student asks for the meaning of a particular word or phrase, then this student could be said to be negotiating meaning (although the negotiation is limited because the learner is not talking directly to his/her interlocutor). In order to be fully communicative, however, this activity is required to have a non-linguistic goal and whether there is such a goal depends on the content of the recording and on what the learners do next. If the content of recording presents information that is true about the world, for example if the speaker describes tourist attractions in London, the activity is potentially communicative. If, however, the speaker in the recording tells information that is purely fictitious, such as voicemail messages which were made up by the textbook author, the learners cannot find out anything about the world. They are not obtaining any useful information; they are just practising language. Communicative tasks require the presentation of real information; fake information sells our learners short.

Simply listening to information about the world is not enough to ensure communicativeness, however. Learners also need to be aware of their non-linguistic goal while they are listening. In other words, learners must know while they are listening to the audio tracks that they will be required to do something with the information they obtain. A follow-up activity could take up one of various forms, such as filling in a table comparing sights in London with those in Tokyo, writing a survey about

London sightseeing spots, interviewing classmates about which sights they would most like to visit, and so on, depending on the time available and the ingenuity of the instructor. So far, we have outlined what a communicative listening task would look like and also what would prevent it from being described as communicative. Readers will probably have noticed that most published listening materials do not fit in with the conditions that we have proposed.

Similar criteria can be applied to written texts. Communicative reading activities also need to have a clear non-linguistic goal: this requires learners to use the information they obtain in the text for some pre-decided purpose. For example, learners could exchange information with classmates, integrate their knowledge and use their synthesis as a basis for a written reflection or presentation. Output activities are not essential, however. Just checking 'True' or 'False' next to a list of facts may be sufficient, if the goal is for learners to build knowledge about a place of interest which they apply in a later part of a communicative task. Skill-building exercises such as skimming or scanning practice aim to have learners copy certain native-speaker habits and therefore do not come close to meeting the conditions set out in VanPatten's definition (this applies equally to explicit pronunciation practice). Reading stories may not involve learning facts about the real world, however, in ordinary life, people read stories for pleasure, so autonomously reading stories for pleasure is undoubtedly a communicative activity.

Since communicative is concerned with meaning and and not with conscious attention to linguistic form, any kind of grammar practice is uncommunicative. Some mechanical grammar drills may not even involve the expression and interpretation of meaning, because learners can complete them without understanding what they are saying or writing. Writing an essay could be communicative if the learner is aware that it is actually going to be read for its content, rather than marked only for accuracy. Writing five sentences to describe what a learner did over the weekend, in order to practice using the past tense, would of course, not be communicative according to VanPatten's (2019)

definition because the goal is simply to practice certain language forms.

Next, let us turn to speaking activities. Activities or tasks which require learners to speak take many forms and many of them cannot be described as communicative if we apply VanPatten's (2019) definition. Reading a dialogue does not require learners to express any of their own meanings. Having learners write a dialogue and then act it out does involve the expression of meaning, but there is no purpose other than to practice language. Having learners ask and answer a set of conversation questions is meaning-focused, but for reasons explained above, in a communicative task learners would probably have to do something with the information they obtained during their interactions with their classmates if there is to be a non-linguistic purpose. Motivated learners may enjoy simply asking and answering questions, and if they genuinely want to ask their interlocutors the questions they are given and are interested in their classmates' answers, then it may be possible to argue that these learners are genuinely communicating. It is our experience, however, that such situations occur very rarely in institutional language classrooms, because teacher and learner interests do not fully coincide.

Role-plays are a speaking activity that have long been favoured by teachers who wish to promote their learners' spoken fluency. They have also been used by teachers of English for Specific Purposes, to try to recreate in the classroom some of the communicative contexts that they are preparing their learners to navigate. Undoubtedly, role-plays allow for the expression and interpretation of meaning and do not necessitate a focus on form. Their non-linguistic purpose is questionable, however. An equally severe flaw is connected to the final element in VanPatten's (2019) definition. Communication always occurs in a given context: this involves the setting, which is the classroom, and the participants, which comprise the learners and the teacher (VanPatten, 2017). This context is always fixed (VanPatten, 2017). Role-plays may attempt to alter the classroom setting, by having learners pretend to be customers or servers in a restaurant, for example, but the

power relationships of the classroom cannot disappear. Learners know that they are interacting with classmates who are merely pretending to be someone else, and they are also aware that their teacher is evaluating them. These factors will influence the way they behave during the role-play. It may be unfortunate and disappointing, but teachers need to accept that it is impossible to recreate 'real-life' situations in the classroom if we wish to teach communicatively. Role-plays, therefore, cannot be described as communicative tasks. This last point may prove controversial, but it is the logical conclusion of the definition of communication we have adopted. Another logical implication is that psycho-social goals are unsuitable for classroom activities. The fixed roles (learners and a teacher) in a classroom profoundly influence the development of its participants' social relationships, and therefore it would be challenging to align learners' individual psycho-social goals with task goals. For this reason, cognitive-informational goals are most suitable for communicative activities.

It may appear that we are arguing for teachers to discard substantial amounts of traditional practice. This is not our goal in this paper. What we are aiming to do is to inform teachers of what can and cannot be described as communicative. It is not our intention to state that non-communicative activities have no place in the L2 classroom, only to argue that several criteria must be met for teaching to be genuinely communicative. In fact, VanPatten (2017) stated that some partially communicative activities, such as those which involve the expression and interpretation of meaning but lack a non-linguistic goal, may have a useful role in that they may help to prepare learners to engage in fully communicative tasks. Skill-building activities like skimming and even pronunciation activities may conceivably have a useful place in the classroom, provided there is data to prove their effectiveness.

Criteria for Evaluating 'Communicativeness'

In the following section, we will present a selection of criteria that have been proposed in the

literature for judging the communicative nature of classroom discourse and materials, discuss their merits and demerits and compare them with VanPatten's (2019) definition of communication and its implications. Finally, there will be a discussion of the value of classifying activities as either communicative or uncommunicative.

Nunan (1987) argued that genuine communication is characterised by five features, which are summarised below:

1. The uneven distribution of information.
2. The negotiation of meaning.
3. Topic nomination and negotiation by more than one speaker.
4. The right of interlocutors to decide whether to contribute to an interaction or not.

(Nunan, 1987)

Item 1 proposes the information gap as a feature of communication, which fits in with VanPatten's (2019) definition to the extent that an information gap can provide an interaction with a non-linguistic goal. However, VanPatten's definition is more comprehensive than Nunan's proposal in that it allows for the fact that not all communication is based around information exchange. The negotiation of meaning is also described more accurately in VanPatten's definition because this occurs only when there is a communication breakdown (VanPatten, 2017). Items 3 and 4, if they are valid claims, present a possible weakness of VanPatten's (2019) definition, which includes nothing about topic negotiation or the rights of interlocutors to be silent. Nunan (1987) states that in genuine communication, "decisions about who says what to whom and when are up for grabs" (p.137). This view of communication is also reflected in item 3 of Fröhlich, Spada, and Allen's (1985) proposed five features of communicative language teaching, which are summarised below:

1. Group work.
2. Focus on meaning.

3. Topic control by the teacher and the students or by the students alone.
4. The use of extended texts.
5. The use of semi-and non-pedagogic materials.

(Fröhlich et al., 1985)

Another possible feature of communication not included in VanPatten's (2019) definition is proposed by Thornbury (1996), who argued that "communicativeness should be defined qualitatively". The quality of communication is high when participants are both frank and considerate and speak willingly on topics that interest both them and their interlocutors (Puchta and Schratz 1993, in Thornbury 1996). It is necessary to consider whether topic nomination and control and 'quality of communication' are useful concepts when considering the meaning of the term 'communicative' language teaching. Firstly, it must be acknowledged that implicit in the views expounded by Nunan (1987), Fröhlich (1985) and Thornbury (1996) is the belief that the term 'communicative' relates primarily to the act of speaking. This is problematic because since communication can involve written and aural texts, we should be able to use the term 'communicative' for teaching of any of the four skills. Even if we focus only on spoken interaction, however, weaknesses in the assertions become apparent. Let us consider a police interrogation. Surely, nobody would deny that this is a communicative event. The person under interrogation would have no control over the topic and would likely not be speaking voluntarily on a topic of interest, so the quality of the communication, according to Puchta and Schratz's (1993, in Thornbury, 1996) description, would be low. These are, therefore, not critical factors in judging the communicative nature of an interaction. One could argue that they are desirable features of conversations between adults who are in an equal power relationship, but desirable qualities are not crucial elements. In addition, it is questionable how relevant they are to the L2 classroom, in which teachers have to make choices about suitable classroom topics and may not be able to leave the content of interactions to the learners. It is also

difficult to see how the proposed criteria could be applied to classes of beginner-level learners who cannot engage in conversations in the L2.

The implicit assumption that communicative activities involve speaking can also be found in other proposals, such as by Pica, Kanagy, and Falodun (1993, in Parry, 2000) and Ellis (1982). Pica et al. classified communicative activities into five types: jigsaw, information-gap, problem-solving, decision-making and opinion-exchange. In line with our arguments earlier in this paper, far more kinds of activity can be imagined. It must also be added that many of the activities suggested by Pica et al. would seem to be suitable for more advanced level learners who possess sa high degree of oral fluency. It may be worthwhile to reiterate here that a strength of VanPatten's (2017) approach is that it is applicable to all teaching contexts, even when the learners are complete beginners, and to teaching which covers all the four skills.

Ellis (1982) proposed five conditions which materials need to meet if they are to be considered communicative. These can be summarised as:

1. The success of an activity generated by the materials is demonstrated by its outcome and not by its process.
2. The speakers must be concerned with what they say, not how they say it.
3. There must be an information gap.
4. Spoken interaction must be negotiated rather than predetermined.
5. Speakers must be allowed to use whatever resources they possess to make meaning, even if this is not typical of native speakers.

Items 1 and 3 fit in with VanPatten's (2017, 2019) assertion that communication has a non-linguistic purpose. Item 2 would seem to explain that when people communicate, they focus on meaning and not on form. This fact is an implicit component of VanPatten's (2019) definition. If people have a non-linguistic purpose when expressing and interpreting meaning, then it is reasonable

to assume that they are not just practising certain language items. The importance of a focus on meaning rather than form constitutes the basis of Littlewood's (2004) proposed continuum of activities that stretches from a non-communicative extreme to 'authentic communication'. The defining feature of the latter is that there is a complete focus on meaning (rather than form) and that the language used is unpredictable. Harmer (1982) also proposed a similar continuum to Littlewood's (2004), and argued that genuinely communicative activities have the following features:

1. Learners have a communicative purpose.
2. Learners have a desire to communicate.
3. Learners are focused on content not form.
4. Learners use a variety of language.
5. There is no teacher intervention when learners make a mistake.
6. The materials do not control the language used by the learners.

Ellis's (1982) view that learners cannot be forced to use certain predetermined language items in communicative activities is echoed in Harmer's items 4 and 6 and is perhaps a reaction against the use of drills in the L2 classroom. While both these proposed continuums may contain some truth, they have limited value because they miss out several of the essential features of communication found in VanPatten's (2019) definition, such as the role of context and the fact that learner output is not required for communication to take place.

Finally, let us consider the argument that distinguishing between communicative and non-communicative activities is unwise because so much depends on teacher implementation. Beaumont and Chang (2011), for example, argued that the dichotomy between traditional and communicative teaching activities is unhelpful, and instead, we should judge practice by learning outcomes. A similar point was made by Harmer (1982) when he claimed that it was problematic to view the term

'communicative' in "an 'all-or-nothing' way: either teaching is communicative or it is not" (p.165). This view, he argued, was a cause of the confusion around the term 'communicative', because, "students can learn to communicate in many different ways and as a result of many different techniques" (Harmer, 1982, p.165). We would disagree. It seems pertinent here to reiterate our point that communicative ability develops through engagement in genuine communication. The argument that any activity that indirectly leads to the development of communicative ability can be called 'communicative' would not only stretch the term beyond its usefulness but also deny Savignon's (2002) core essence of CLT. Beaumont and Chang (2011) contend that "(r)eading aloud and dictation can be rendered communicative in any number of ways" (p.298). Applying VanPatten's (2019) definition helps us to examine this claim. When reading aloud or doing a dictation, learners are not expressing their own meaning, and there is no discernible purpose except to practice the language, so it is impossible to call these activities 'communicative'. Some teachers may be unhappy with such a declaration because it may be construed as a criticism of some of their cherished practices, but it is a logical conclusion, nonetheless.

Conclusion

Harmer (1982) argued that the term 'communicative' should not be applied to methodology for two reasons. Firstly, because "it will prohibit the use of many tried and tested techniques" (p.165). We believe, however, that if traditional practice is shown to be uncommunicative, practitioners need to be willing to acknowledge that and duly revise their perspective. We need to reconsider the value of what has come to be viewed as 'tried and tested' and remind ourselves of what CLT was first intended to be. We concur with Thornbury (1998), who argued that CLT, as it was originally conceived, has had little impact on current language teaching practice, and this has to change if the field is to move forward. Secondly, Harmer (1982) argued that the term 'communicative' "will have

to have a definition so broad as to be meaningless" (p.165). We disagree. The definition of communication proposed by VanPatten (2019), is, we would argue, actually a narrow one because it lays down conditions that are so precise. Indeed, teachers who adopt this definition to inform and guide their practice may feel constrained because it deems much of current L2 teaching practices uncommunicative. It may well be impossible to reconcile CLT with traditional language teaching, despite the arguments of some (e.g. Pica, 2000). However, VanPatten's (2017) perspective offers a radical, clear-sighted and judicious roadmap to move CLT forward.

In this paper, we have given our support to the position advocated by VanPatten (2017) that communicative language teaching needs to be informed by a definition of communication. The definition proposed by VanPatten (2019) is comprehensive yet succinct. With the knowledge it gives, practitioners can analyse teaching practices and materials to judge whether they can be called 'communicative': this is to say, whether they offer opportunities for genuine communication. Such judgments are significant because learners learn to communicate by communicating and teachers need to organise their classes so that their learners engage in communicative events. This is not to imply that we advocate only the use of communicative activities in class: non-communicative activities may have a role to play. What we are arguing in this paper is that it is crucial for teachers to know the difference between communicative and uncommunicative practices according to VanPatten's definition of communication.

VanPatten's (2017) position is, as far as we know, unique in the literature. We have reviewed other attempts to analyse the communicative nature of classroom discourse and teaching materials, but they all miss out on some of the essential features of communication, such as the importance of context, and frequently apply only to speaking activities. We believe that a clear distinction between communicative and uncommunicative teaching is a necessary consequence of adopting VanPatten's (2019) definition, and this is to be welcomed. This distinction can help practitioners to revise their

practices in a way that is logically consistent and illuminates the path towards a different type of instruction that fits in with the true aims of CLT that were first proposed more than four decades ago.

References

Arnold, J., Dornyei, Z., & Pugliese, C. (2015). *The principled communicative approach.* London: Helbling.

Beaumont, M., & Chang, K. S. (2011). Challenging the traditional/communicative dichotomy. *ELT Journal*, *65*(3), 291-299.

Celce-murcia, M., Dörnyei, Z., & Thurrell, S. (1997). Direct approaches in L2 instruction: A turning point in communicative language teaching?. *TESOL Quarterly*, *31*(1), 141-152.

Crookes, G., & Gass, S. M. (eds.) (1993). *Tasks and language learning: Integrating theory and practice.* Clevedon: Multilingual Matters.

Ellis, R. (1982). Informal and formal approaches to communicative language teaching. *ELT Journal*, *36*(2), 73-81.

Fröhlich, M., Spada, N., & Allen, P. (1985). Differences in the communicative orientation of L2 classrooms. *TESOL Quarterly*, *19*(1), 27-57.

Harmer, J. (1982). What is communicative? *ELT Journal*, *36*(3), 164-168.

Hinkel, E. (Ed), *Handbook of research in second language teaching and learning* (Vol. II, pp. 541-557). New York: Routledge

Littlewood, W. (2004). The task-based approach: some questions and suggestions. *ELT Journal*, 58(4), 319–326.

Littlewood, W. (2011). Communicative language teaching: An expanding concept for a changing world, In E. Hinkel (Ed), *Handbook of research in second language teaching and Learning* (Vol. II, pp. 541-557). New York: Routledge

Nunan, D. (1987). Communicative language teaching: Making it work. *ELT Journal*, *41*(2), 136-145.

Parry, M. (2000) How communicative are introductory undergraduate level Japanese language textbooks? *Japanese Studies*, 20(1), 89-101.

Pica, T. (2000). Tradition and transition in English language teaching methodology. *System*, 28(1), 1-18.

Pica, T., Kanagy, R., & Falodun, J. (1993). Choosing and using communication tasks for second language instruction. *Multilingual Matters*, 9-9., in G. Crookes & S. Gass (Eds), *Tasks and language learning: Integrating theory and practice* (Clevedon: Multilingual Matters, 1993), pp. 9-34.

Puchta, H., and Schratz, M. (1993). T*eaching teenagers: Model activity sequences for humanistic language learning*. Harlow: Longman.

Sakui, K. (2004). Wearing two pairs of shoes: Language teaching in Japan. *ELT Journal*, *58*(2), 155-163.

Savignon, S. (1997). *Communicative competence theory and classroom practice: Texts and contexts in second language learning (2nd edition)*. New York, NY: McGraw-Hill.

Savignon, S. J. (2002). Communicative curriculum design for the 21st century. In *English Teaching Forum* (Vol. 40, No. 1, pp. 2-7). Retrieved from https://americanenglish.state.gov/files/ae/resource_files/02-40-1-c.pdf

Thornbury, S. (1996). Paying lip-service to CLT. *EA Journal*, *14*(1), 51-63.

Thornbury, S. (1998). Comments on Marianne Celce-Murcia, Zoltán Dörnyei, and Sarah Thurrell's "Direct approaches in L2 instruction: A turning point in communicative language teaching?": A reader reacts…. *TESOL Quarterly*, *32*(1), 109-116.

VanPatten, B. (2017). *While we're on the topic: BVP on language, acquisition, and*

classroom practice. Alexandria, VA: The American Council on the Teaching of Foreign Languages

VanPatten, B. (2019, May 23rd). Tenets of comprehension-based language teaching with a focus on communicative ability [PDF File]. Retrieved from https://www.inputandmore.com/talkin-l2-resources

Bridging Learning and Testing in an EFL Curriculum: Pursuing the Effective Use of a Commercially Produced Speaking Test

Kazuharu Owada

Ritsumeikan University

Yuko Shimizu

Ritsumeikan University

Abstract

As part of the English-language curriculum for a new language program with non-English majors at a private university, we have adopted a computer test that emphasizes the development of speaking skills. It was not realistic to develop a localized speaking test because of the limited resources of our institute. Therefore, we decided to adopt a commercially available speaking test. In this study, we conducted a questionnaire survey of the students immediately after they took the speaking tests both at the beginning and at the end of the program to observe the validity of the test use from the learners' perspective. We briefly describe quantitative results for both the speaking test and the questionnaire. In addition, we present the results of text analysis of the post-questionnaire by using KH Coder (Higuchi, 2016, 2017). On the basis of the results we have obtained in this study, we conclude by noting the importance of the "alignment" between assessment and content in order to make the program successful and to provide visible results to the learners.

Keywords: EFL curriculum, computer-delivered speaking test, questionnaire, text analysis

Introduction

Lado (1961) once identified the ability to speak as a "most highly prized skill yet the least developed and least practiced in the language testing field" (p. 299). It is true that a speaking test is difficult to put into practice and the subjectivity of evaluation will cause a problem of reliability. Systematic testing of production skills is hard to realize particularly in a classroom situation, and the ability to speak does not have a strong impact on the learners' final grades, as Mewald et al. (2013) point out in the context of foreign-language education in Austria. This might apply to many English classrooms in Japan. However, the recent trend of English education in Japan shows a compelling need for finding a way or ways to assess speaking aspects in learning. To be more exact, arguments on the adoption of private-sector English tests for university entrance examinations somewhat highlighted speaking tests. For example, on November 1, 2019, the Japanese Ministry of Education, Culture, Sports, Science and Technology (MEXT) postponed the introduction of such tests used as part of university entrance examinations for academic year 2020.

As Watanabe (2016) points out, education will not be changed simply by changing what tests are used. In the process of teaching and learning, tests or assessments as a system should be used to increase students' learning. No teachers will deny that the foundational principle underlying the course of study in Japan is to promote students' ability to communicate as part of English education. It will be no small challenge to find out how to ensure the relationship of teaching and testing and also how performance skills are measured and evaluated.

As a small-scale practice, we introduced a speaking test in our English program to connect learners and curriculum, as well as an assessment. In this study, we used one of the commercially available speaking tests in our English program with non-English majors. By analyzing the results of the questionnaire conducted after the speaking tests, we examined the validity of the test use from the test-takers' perspective in particular.

Literature Review

The International Association for the Evaluation of Educational Achievement (IEA) based in Amsterdam conducts various large-scale comparative studies of education. The backbone of their research is to pursue the relationship among *intended curriculum*, *implemented curriculum*, and *achieved curriculum*. In the field of testing and assessment, validity is one of the important factors. It is "the degree to which scores on an appropriately administered instrument support inferences about variation in the characteristic that the instrument was developed to measure" (Cizek, 2012, p. 35). Validity is fundamental in evaluation of tests in use. In the field of educational measurement, a relatively new concept, *alignment*, is often referred to in discussions of validity. By citing several researchers, Case et al. (2004) define it as "the degree to which the components of an education system—such as standards, curricula, assessments, and instruction—work together to achieve desired goals" (p. 2).

That matches the concept of IEA. It is crucial to connect objectives, activities, and assessment for successful learning. We followed such a concept when we designed our English program, and we decided to validate the use of a test in our English program. That is, we introduced a commercially available speaking test to measure our students' learning since fostering communication skills is required as part of the program. Cizek et al. (2018) argue that the form of a test must go with its intended content to show evidence of validity. The match is "key not only to test construction and accurate score interpretations, but also bears on the justification for using test scores in any intended way" (p. 507). Such an aspect of validity evidence is usually referred to as evidence based on test content in the Standards for Educational and Psychological Testing (American Educational Research Association et al., 2014, p 14).

Why Computer-Delivered Assessment?

Generally, a speaking test, which involves face-to-face interaction, is considered as one of the direct tests. It is true, however, that direct speaking tests are impractical when we test a large number

of test-takers. This practicality problem may be one of the reasons why classroom teachers are reluctant to incorporate speaking tests into assessment.

Fast-paced changes in technology influence many aspects in education, including the assessment system. These days, some speaking tests require the test-takers to speak into a microphone using a computer, telephone, or smartphone. Such semi-direct, or person-to-machine, tests are not equivalent to direct speaking tests. However, some studies collected evidence to show that high correlations can be achieved fairly systematically on well-designed indirect tests of L2 speaking (Fulture, 2003), which may satisfy test administrators and teachers when they decide to use such assessments in their programs.

Test-takers' reactions are important when validating a test. Yonezaki (2016), based on five studies conducted between 1993 and 2015, summarizes test-takers' reactions to testing formats that involve direct and indirect speaking. According to her analysis, two studies (1992, 1994) reported in O'Loughlin (2001) and a study by Nakamura (2015) indicate test-takers' preference for direct testing formats, while a study by Brown (1993) indicates a stronger preference for semi-direct speaking tests. However, Qian (2009) reported that test-takers had no particular preferences. These inconsistent results suggest that there is no one right decision in the testing context. Thus, we must validate the test use in our program according to our educational situation.

In general, computer-delivered speaking tests are said to have some advantages for the administrative side: standardized testing and efficient test procedure (Zhou, 2015). Those advantages were among the reasons we adopted a commercially produced speaking test called Versant. However, we keenly feel the need to validate our decision. Therefore, the present research serves as the first step of the validation process.

Why Versant?

We had to carefully consider what assessment tool to use. Commercially available tests have already proved their reliability in many cases. Therefore, as test users, we had to examine the validity

and practicality of the test for use in our situation. Fortunately, we had experiences of using one such commercially available speaking test called Versant with students who participated in short-term study-abroad programs (Shimizu et al., 2014).

Versant was developed by Pearson Inc. and is currently marketed in Japan by Nihon Keizai Shimbun Inc. Therefore, we understood the test construct and plausibility of the test with Japanese university students. In addition, a study on PhonePass, which is the first version of Versant, demonstrated a significant level of concurrent validity with respect to the TSE exam, r = .88 (Townshed & Todic, 1998, Pearson Education Inc., 2019).

Because of the introduction of several modifications, a number of additional validation studies have been conducted. With each modification, the accuracy of the test has improved, but the scores still correlated highly with previous versions. In addition, construct representation and concurrent validity evidence of automated speaking tests support the score interpretation and use (Bernstein, et al., 2010). Therefore, Versant can help test-takers speak English, as the test-takers' active participation is encouraged in each test task (see Table 1).

Table 1

Test Tasks in Versant

Part	Task	# of items	Interactional relationship
A: Reading	Read the eight sentences one at a time.	8	Non-Interactional
B: Repeats	Repeat sentences that they hear verbatim.	16	Non-Interactional
C: Short Answer Question	Listen to spoken questions and answer each question with a single word or short phrase.	24	Interactional: One-way
D: Sentence Builds	Listen to three short phrases and rearrange them to make a sentence.	10	Non-Interactional
E: Story Retelling	Listen to a brief story and describe what happened in their own words.	3	Non-Interactional
F: Open Questions	Listen to spoken questions that elicit an opinion, and provide an answer with an explanation.	2	Non-Interactional

Background to the Study

Introduction to the College of Gastronomy Management and its Language Curriculum

The College of Gastronomy Management started in April 2018 at Ritsumeikan University. On the basis of the concept of economics and management, the college aims to rear a new generation of food industry leaders by connecting humanity and the scientific aspects of gastronomy. All students are required to obtain at least 124 credits to graduate. Among them, eight credits are allotted to the English courses; six are allotted to additional foreign-language classes. The purpose of foreign-language education in this college is to build students' understanding about not only foreign language and culture but also society, history, geography, and industry. Through foreign-language education, the college hopes to help spread information about Japanese gastronomic culture. Additionally, students will develop collaboration and communication skills for problem solving in a multicultural world. With the above-mentioned purpose in mind, we designed our English curriculum.

The English program was designed to follow the three stages of learning: Step 1 (first and second semesters) to train students to be autonomous learners, Step 2 (third semester) to train students to be career-minded learners, and Step 3 (fourth semester or later) to train students to be expert learners. The first two steps are taught by teachers of English. In Step 1, all the first-year students are to take three English courses in both spring and fall semesters (see Table 2). Each course is separately taught and evaluated by different instructors; however, the content is integrated to some degree. As in Table 2, the courses provided in Step 1 are extremely important for preparing the students for the courses in Step 2, the English Workshop in particular. Therefore, students are encouraged to use English as much as possible throughout the tasks and class work provided for each course.

The aim of Study Skills α1 and α2 is to develop students' reading and listening skills to allow them to function effectively in an academic environment. In addition to basic study skills, first-year students will develop learning attitudes, which are appropriate for university study, including time management, note taking, and the use of graphic organizers.

The aim of Study Skills β1 and β2 is to develop students' speaking and writing skills to allow them to function effectively in an academic environment. Through activities including discussion and presentations, students will develop their communication skills to deliver messages to others. Students will learn basic paragraph organization skills and practice summary writing and essay writing based on topics related to Gastronomy Management.

Activities in CALL1 and CALL2 will include practice in pronunciation, reading aloud, and shadowing to enhance the intelligibility of students' pronunciation. With the computer as a tool, students will learn basic skills such as typing and making Word and PowerPoint files required in Study Skills β1 and β2.

Since the college accepts diverse students in terms of their English ability, first-year students are placed into the four levels based on the results of the English placement test conducted at the very beginning of April. The four levels are Advanced (AD), Upper-Intermediate (UI), Intermediate (IM), and Pre-Intermediate (PI).

Table 2

Structure of the English Program and Skills to Learn

<table>
<tr><th colspan="2" rowspan="2">Skills learned in English Program</th><th rowspan="2">Areas in CEFR</th><th colspan="2">Step 1</th><th>Step 2</th></tr>
<tr><th>First semester</th><th>Second semester</th><th>Third semester</th></tr>
<tr><td colspan="2">Reading skills</td><td>UNDERSTANDING</td><td rowspan="3">Study Skills α1</td><td rowspan="3">Study Skills α2</td><td rowspan="2">English for Career Development</td></tr>
<tr><td colspan="2">Listening skills (incl. lecture listening skills)</td><td>UNDERSTANDING</td></tr>
<tr><td colspan="2">University-level study skills</td><td>UNDERSTANDING</td><td>--</td></tr>
<tr><td rowspan="2">Presentation skills</td><td>Academic written presentation skills</td><td>WRITING</td><td rowspan="2">Study Skills β1</td><td rowspan="2">Study Skills β2</td><td rowspan="5">English Workshop</td></tr>
<tr><td>Academic oral presentation skills</td><td>SPEAKING</td></tr>
<tr><td colspan="2">Seminar skills (e.g., practice in academic listening and speaking, giving opinions, presenting an argument, holding the floor, bringing in other speakers, chairing a mini-conference, facilitating a group)</td><td rowspan="2">SPEAKING & WRITING</td><td rowspan="2">--</td><td rowspan="2">--</td></tr>
<tr><td colspan="2">Cross-cultural communication skills (incl. cultural awareness)</td></tr>
<tr><td colspan="2">Research skills (incl. digital research skills)</td><td>UNDERSTANDING</td><td rowspan="2">CALL 1</td><td rowspan="2">CALL 2</td></tr>
<tr><td colspan="2">ICT literacy skills (e.g., e-learning skills, pronunciation skills, ICT-based presentation skills)</td><td>SPEAKING $ WRITING</td></tr>
</table>

Tests Adopted in our Program

Since the College of Gastronomy Management was a completely new college, there was no way of identifying and/or simulating the English level of newly enrolled students when we designed our program. We, as curriculum designers, assumed that we had to prepare a curriculum to meet the needs of various learners in terms of their English proficiency levels. During the orientation program for first-year students, the students take an English placement test. We use the Computerized Assessment System for English Communication (hereinafter CASEC), a computer-adaptive test originally developed by the Eiken Foundation of Japan and currently developed and operated by The Japan Institute for Educational Measurement, Inc.

Table 3

Tests Used in the English Program

At the time of enrollment	Spring Semester (April)	Fall Semester (December)
CASEC (Placement Test)		CASEC
	Versant	Versant
	Grammar Test	Grammar Test
	Vocabulary Test	Vocabulary Test

In order to provide visible results of the learning to our students, several measurement tools including two external tests, CASEC and Versant, are used (See Table 3). As a placement test, students are to take CASEC in April before the semester starts, as mentioned above. Then, by using some class time, the students take a Versant speaking test, which is the main concern in our current research. In addition, they take grammar and vocabulary tests for diagnostic purposes. Those four tests are also conducted at the end of the fall semester to observe progress and change over time in their English ability.

Method

The purposes of this study are: first, to create a questionnaire survey for Japanese EFL

learners and analyze the survey data to understand learners' attitudes toward a commercially available online computer-delivered speaking test, and second, to examine the test's face validity—that the test measures what it appears to measure—to determine whether students consider this test useful and meaningful.

Materials and Procedure

As the first step, the students took the first Versant test and answered the pre-questionnaire asking for their feedback on that test-taking experience in April. As the second step, they took the second Versant test and answered the post-questionnaire asking for their feedback on that test-taking experience in December. We conducted a statistical analysis of Versant scores and pre- and post-questionnaires as the third step. In addition, we conducted text analysis of the post-questionnaire by using KH Coder (Higuchi, 2016, 2017).

Participants

Participants in the study were non-English-major students who had entered the university in April 2018. Out of 355 students, 328 students, whose first language was Japanese, took the speaking tests. All the first-year students took CASEC as a placement test in April. Then after being divided into four levels, they started taking three English classes per week. The mean score of CASEC was 527.51 (out of 1,000 points) (*SD* = 105.01) in April and 542.03 (*SD* = 97.55) in December.

Versant was conducted in April and December as part of the program assessment. After each time, we asked the participants to answer the questionnaires, in other words, pre- and post-questionnaires, to learn about their attitudes toward the test-taking experience. In other words, in April the participants took the Versant test and right after it answered the pre-questionnaire; in December, they took the Versant test again and answered the post-questionnaire. Out of the 328 students, 272 students took the Versant test twice. Then out of those 272 students, 271 answered both pre- and post-questionnaires.

Pre- and Post-Questionnaire After the Versant Test

We developed a questionnaire used in this study based on previous studies such as Xi et al. (2016); Fan (2014); and Zhou and Yoshitomi (2019). The pre-questionnaire after the first Versant test in April included 13 closed-ended items (5-point Likert scale) for descriptive statistical analysis and three open-ended items for text analysis using KH Coder. On the other hand, the post-questionnaire after the second Versant test in December included the same items as the pre-questionnaire and one additional closed-ended item. Although all the questionnaire items were written in Japanese, the first language of the participants, they were all translated into English here for the sake of discussion.

As for the scale of each item, a 5-point Likert scale (1 = *strongly disagree*, 2 = *disagree*, 3 = *neither disagree nor agree*, 4 = *agree*, 5 = *strongly agree*) was used (See Appendix A). The open-ended items were as follows:

A. *How did you feel about taking this test?*

B. *How do you feel about taking this test as part of an English curriculum at the university?*

C. *What kind of English studies do you think you need to get high scores in this test?*

To understand such unstructured text data obtained by these three items, we decided to conduct text analysis using KH Coder, which is an open-source software for qualitative data analysis.

Results

Versant and Closed Items in the Pre- and Post-Questionnaires

The results of the speaking test and closed items conducted in April and December were compared and analyzed (Shimizu & Owada, 2019). Therefore, here we briefly summarize the results.

Regarding the scores of Versant, which range from 20 to 80, the mean score in April was 33.86 (SD = 5.13), and it was 36.39 (SD = 5.55) in December. The difference in the mean score of 2.53 indicated a significant improvement based on the paired t-test, $t(327) = 2.53$, $p < .001$, $d = .74$. As to

the closed items, a translated (into English) version of the questions is shown in Appendix A. Since items 3 and 4 were negatively worded, we reversed the scores before the analysis. To see the difference between the first test-taking experience and the second one, paired *t*-tests were conducted, and, as a result, significant differences were observed in seven items. Broadly speaking, students' attitudes toward the computer-delivered speaking tests improved in December compared to April. As students seemed to believe that the test properly measures their speaking ability, we can tentatively conclude that the students showed positive attitudes toward this measurement tool for speaking. The results indicated that they consciously tried to pronounce more clearly and to speak more fluently, when taking the tests. We hope that the introduction of the speaking test in the curriculum will positively affect the students' learning. Some results indicated that students found a connection between what they learned in class and the tasks in the speaking test.

Although the quantitative analysis showed fairly favorable results from the test-takers, there is a possibility that their honest opinions might not have been fully revealed in five-point scales. Therefore, we sought to discern their honest opinions through the open-ended items.

Text Analysis of Open-Ended Items in the Pre- and Post-Questionnaires

For qualitative analysis, we focus on the post-questionnaire in this study. As for the open-ended items, we used KH Coder. First, we made a list of frequently used words to obtain an overall picture. Then, we conducted a co-occurrence analysis to see how each word is connected to another. Co-occurrence refers to the presence of two words in a sentence. We were interested to find out what kind of same word pairs co-occur frequently in student comments. For the co-occurrence, KH Coder uses the Jaccard coefficient.[ii] The Jaccard coefficient ranges between 0 and 1. A score of 0 is obtained when two words are completely different, and a score of 1 when two words are identical. A score of 0.1 means weak correlation between the two words, and 0.3 means strong correlation.

A total of 271 students responded to the open-ended items in Japanese. In item A, 177 out of 271 students provided their answers (valid response rate: 65.3%); in item B, 168 out of 271 students

provided their answers (valid response rate: 61.9%); and in item C, 170 out of 271 students provided their answers (valid response rate: 62.7%). The results of word frequencies and co-occurrence network analysis are discussed below item by item. All the responses quoted from the students have been translated into English.

Item A. Figure 1 shows the frequent list of words for A. *How did you feel about taking this test?* The word "difficult" was the most frequent with 77 occurrences. As for the network shown in Figure 1, we can see some clusters of words along with the similar Jaccard coefficients. A cluster at the top left indicates that the students found it "difficult to remember the question in each part."

Table 4

Frequency List of Words in Item A

Rank	Words	Freq.	Rank	Words	Freq.
1	difficult	77	11	spring	12
2	think	32	12	test	11
3	myself	22	13	say	11
4	English	20	14	nervous	10
5	feel	18	15	voice	10
6	able to listen	18	16	previous	9
7	take	16	17	listen	9
8	speak	16	18	problems	9
9	can	14	19	time	7
10	sound	12	20	class	7

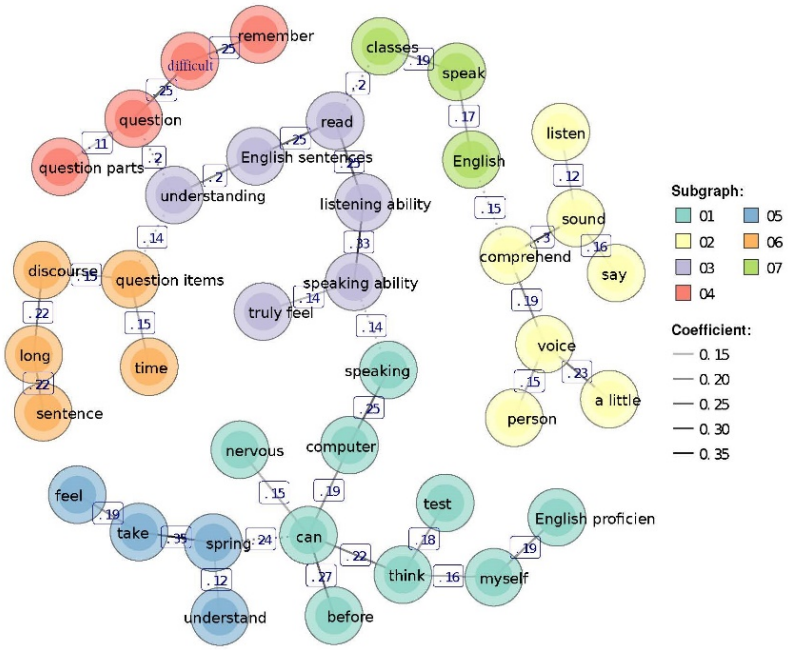


Figure 1. Co-occurrence Network in Item A.

The students wrote a variety of responses to item A, which provides us with valuable information on how they felt about the test-taking experience. Here are some of the examples in terms of positive and negative comments.

[Positive comments]

This time I was able to take the test with relative ease.

[Negative comments]

It's difficult to speak right after memorizing the sentences.

It was difficult to construct sentences by ordering the words.

I hesitated to start speaking while nobody was speaking.

I forget what is talked about when listening to a long discourse.

Just listening to the sound, with nothing to look at, was difficult.

The speed of the voice was so fast. I want the test voice to pronounce words more clearly.

Item B. The results are presented below for B, *How do you feel about taking this test as part of an English curriculum at the university?* As in Table 5, positive words such as "good" and "opportunity" were observed.

Table 5

Frequency List of Words in Item B

Rank	Word	Freq.	Rank	Word	Freq.
1	think	130	11	understand	11
2	good	60	12	university	10
3	myself	25	13	conduct	8
4	test	21	14	speaking ability	7
5	take	16	15	can	7
6	English	15	16	take	6
7	especially	13	17	class	6
8	opportunity	12	18	English ability	5
9	speak	12	19	feel	5
10	speaking	11	20	scoring	5

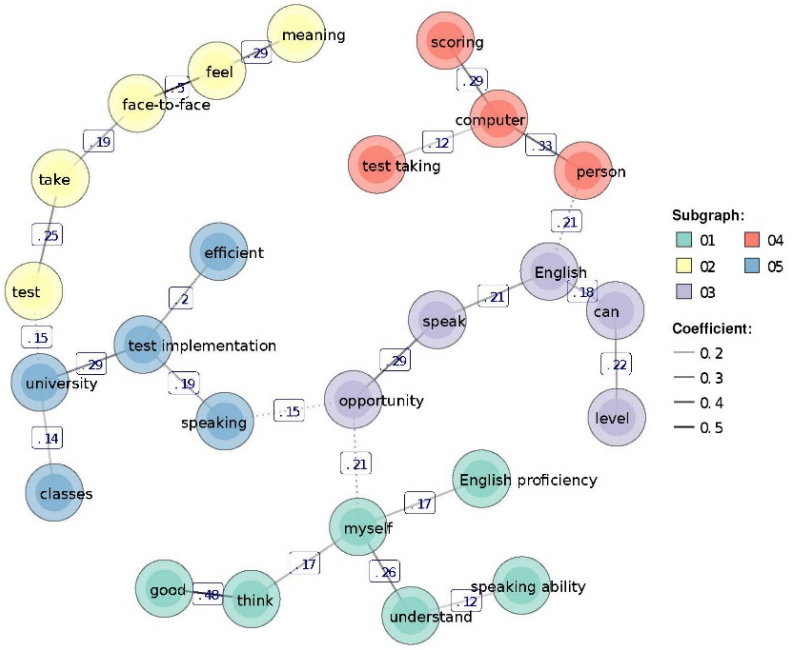


Figure 2. Co-occurrence Network in Item B.

From Figure 2, we can see that the students seemed to "think it is good to understand their speaking

ability" and that "it is an opportunity to speak and know their level." Other positive comments and negative comments are shown below:

[Positive comments]

This test is efficient timewise, and I feel less nervous than face-to-face.

My English is evaluated from different perspectives, so I can understand my level.

[Negative comments]

This test is good for students who are good at English, but not so for those whose level is average or below.

I don't know about the criteria of this test, so I have some reservations.

Item C. The results are presented below for C, *What kind of English studies do you think you need to get high scores in this test?* As in Table 6, since "English," "speak," "pronunciation," "regularly," and "listen" were in the top-6, the students seemed to think it was necessary to speak and listen to English regularly to improve pronunciation. This can also be observed in the cluster of words in Figure 3.

Table 6

Frequency List of Words in Item C

Rank	Word	Freq.	Rank	Word	Freq.
1	English	70	11	speaking	11
2	think	38	12	tackle	10
3	speak	35	13	every day	10
4	pronunciation	25	14	necessary	10
5	regularly	24	15	learning	9
6	listen	22	16	class	9
7	EnglishCentral	19	17	person	9
8	listening	19	18	actively	9
9	practice	15	19	opportunity	8
10	native speakers	13	20	myself	8

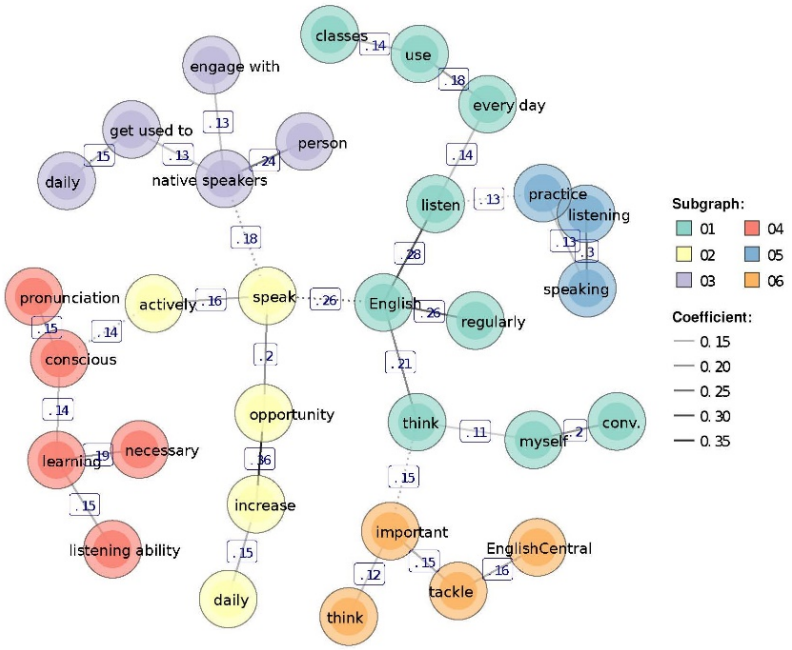


Figure 3. Co-occurrence Network in Item C.

From some excerpts of the actual student responses below, we can see that the students seemed to think it was important to do the following:

- try to speak to a computer as if they were speaking to a person,
- think and speak on the spot rather than speaking from notes,
- take every opportunity to listen to English spoken by English native speakers,
- pay attention to pronunciation and grammar and also do e-learning (*English Central*),
- speak loudly first of all and also pronounce words smoothly,
- use English on a daily basis,
- participate actively in English classes and self-access learning centers.

It is worth noting that these responses reflect the positive relationship between learning and testing.

Discussion and Conclusion

When designing a new English program, we ensure that it links curriculum, learners, and

assessment. As one of the assessment tools, we used a computer-delivered speaking test called Versant. Then we were able to observe a significant improvement in the students' speaking abilities by comparing the scores of Versant in April and December.

Ideally speaking, a local test should be developed to connect our program and learners. In this light, the assessment tool we adopted was not a test specifically designed for our program. Therefore, we felt the need to examine the validity of test use from the perspective of both learners and the program by conducting a questionnaire to examine their attitudes toward the test-taking experience.

So far, through our text analysis, we have seen that although some students found the instructions of the Versant test a little difficult to understand, they did not find it stressful to speak to a computer in that test. The Versant test served not only as a measurement tool for us but also as a good opportunity for the students to understand their level of English. In fact, some students mentioned that taking the Versant test was an opportunity to know their English level. It is also worth mentioning that they have their own ideas about how best to study English in order to obtain high scores on the Versant test, for example, doing e-learning and active participation in regular English classes and self-access learning centers. Therefore, we can safely say that it was a good choice to use Versant in our program, given the favorable student feedback.

One thing we must keep in mind is that in its speaking construct Versant does not provide any information about some aspects of the students' speaking skills that are not measured in this test. Versant does not include all aspects of speaking skills in its speaking construct. Although such skills as turn-taking strategies, facilitating, and negotiating are included in our program, they are not included in the construct of the Versant test. In other words, unlike the speaking construct in the Versant test, those skills are incorporated in our curriculum and measured and evaluated in the form of speaking tasks as part of the students' learning process in a classroom. Therefore, we can say that it is not only that teaching and learning support each other, but that testing and learning also support each other in our program.

Finally, two limitations of this study need to be considered. First, since we only analyzed the data from the first-year students in academic year 2018, we need to accumulate more data and conduct a more in-depth analysis of the students. As the College of Gastronomy Management is a new college at our university, we are not sure yet how diverse our future students will be in terms of their English proficiency. We currently offer four levels of classes in our English program for students with diverse levels of English and backgrounds. Therefore, we will continue to use the Versant test in our English program so that we can accumulate more data. By analyzing the data longitudinally to capture our students' English proficiency, we are hoping to provide a fair assessment for our students.

As a second limitation, we did not take into account the level of English proficiency of the students in our text analysis. It is most certainly the case that students have different attitudes toward the same measurement tool, in other words, Versant, depending on their English proficiency levels. Therefore, we need to investigate whether there is any difference in terms of student responses to the questionnaires between students of different English levels.

Notes

For more information on Jaccard coefficient, go to http://koichi.nihon.to/cgi-bin/bbs_khn/khcf.cgi?no=1313&mode=allread.

Acknowledgements

This work was supported by Grants-in-Aid for Scientific Research (Grant # JP18K00893)

References

American Educational Research Association, American Psychological Association, National

Council on Measurement in Education, Joint Committee on Standards for Educational and Psychological Testing. (2014). *Standards for educational and psychological testing*. Washington, DC: AERA.

Bernstein, J., Moere, A.V., & Cheng, J. (2010). Validating automated speaking tests. *Language Testing, 27*(3), 355-377.

Brown, A. (1993). The role of test-taker feedback in the development process: Test taker's reactions to a tape-meditated test of proficiency in spoken Japanese. *Language Testing*, *10*, 277-304.

Case, B.J., Jorgensen, M.A., & Zucker, S. (2004). *Alignment in educational assessment*. New Jersey: Pearson Education.

Cizek, G.J. (2012). Defining and distinguishing validity: Interpretations of score meaning and justifications of test use. *Psychological Methods*, *17*(1), 31–43.

Cizek , G.J., Kosh, A. E., & Toutkoushian, E.K. (2018). Gathering and evaluating validity evidence: The generalized assessment alignment tool. *Journal of Educational Measurement*, *55*(4), 477–512.

Fan, J. (2014). Chinese test takers' attitudes towards the Versant English Test: A mixed-methods approach, *4*(6), *Language Testing in Asi*a. doi.org/10.1186/s40468-014-0006-9

Fulture, G. (2003). *Testing second language speaking*. London: Routledge.

Higuchi, K. (2016). A two-step approach to quantitative content analysis: KH Coder tutorial using Anne of Green Gables (Part I). *Ritsumeikan Social Sciences Review*, *52*(3), 77-91.

Higuchi, K. (2017). A two-step approach to quantitative content analysis: KH Coder tutorial using Anne of Green Gables (Part II). *Rtsumeikan Social Sciences Review*, *3*(1), 137-147.

Lado, R. (1961). *Language testing: The construction and use of foreign language tests*. London: Longman.

Mewald, C., Gassner, O., Rainer, B., Lackenbauer, F., & Siller, K. (2013). Testing speaking for the E8 standards. *Technical Report 2012*. Salzburg: BIFIE.

Nakamura, K. (2015). Comparability study of PBT and CBT speaking test of English for EFL learners. *The JACET 54th (2015) International Convention Book*. 41-42.

O'Loughlin, K. J. (2001). *The equivalence of direct and semi-direct speaking tests*. Cambridge, UK: Cambridge University Press.

Pearson Education, Inc. (2019). *Versant™ English Test: Test description and validation summary*. Retrieved from https://www.pearson.com/content/dam/one-dot-com/one-dot-com/english/SupportingDocs/Versant/ValidationSummary/Versant-English-Test-Description-Validation-Report.pdf

Qian, D.D. (2009). Comparing direct and semi-direct modes for speaking assessment: Affective effects on test-takers. *Language Assessment Quarterly*, *6*, 113-125.

Shimizu, Y., Kirimura, R., & Nozawa, T. (2014). Keizaigakubu eigo tanki ryugaku puroguramu ni okeru supikingu tesuto no jisshi to sono kekkahokoku [Effects of short-term study abroad programs for Economics students: Based on analysis of speaking tests]. *Ritsumeikan Kotokyoiku Kenkyu*. *14*, 91-102.

Shimizu, Y., & Owada, K. (2019). Daigaku eigo karikyuramu ni okeru gaibu supikingu tesuto no donyu to shitsumonshi ni yoru datosei kensho [A validation study of a commercially produced speaking test in a university-level English language program]. *Studies in Language Science Working Papers*, *9*, 1-14.

Townshed, B., & Todic, O. (1998). *Comparison of PhonePass™ Testing with the Educational Testing Service® Test of Spoken English™ (TSE®)*. ORDINATE Corporation.

Watanabe, Y. (2016). *Jyuken shido wa waruku nai* [Teaching to the entrance exam is not bad]. *ELPA Vision*. No. 2, 11-12. Association for English Language Proficiency Assessment.

Xi X., Schmidgall J., & Wang Y. (2016). Chinese users' perceptions of the use of automated scoring for a speaking practice test. In G. Yu & Y. Jin (Eds.), *Assessing Chinese Learners of English* (pp. 150-175). London: Palgrave Macmillan.

Yonezaki, M. (2016). A comparative analysis of semi-direct speaking testing and direct speaking

testing for Japanese EFL learners. *International Journal of Curriculum Development and Practice*, *18*(1), 27-38.

Zhou, Y. (2015). Computer-delivered or face-to-face: effects of delivery mode on the testing of second language speaking. *Language Testing in Asia*, *5*(2). doi.org/10.1186/s40468-014-0012-y

Zhou, Y., & Yoshitomi, A. (2019). Test-taker perception of and test performance on computer-delivered speaking tests: the mediational role of test-taking motivation. *Language Testing in Asia*, *9*(10). doi.org/10.1186/s40468-019-0086-7

Appendix A: Questionnaire Items with Descriptive Statistics and Results of *t*-test (*N* = 233)

		April *M* (*SD*)	December *M* (*SD*)	Difference (Dec-Apr)	*t*-value	*p*-value
1	I was able to take this test without any problems.	2.43 (1.31)	3.14 (1.23)	0.71	-7.260	< .001
2	The directions of the test were easy to understand.	2.79 (1.12)	3.47 (1.16)	0.68	-7.175	< .001
3	I had some reservations about talking to a computer through a mic. (Reversed)	2.94 (1.36)	2.78 (1.26)	-0.15	1.501	.187
4	I felt nervous about taking this test in comparison with a face-to-face interview test. (Reversed)	2.06 (1.18)	1.95 (1.07)	-0.12	1.324	.187
5	I need to study hard in an English class in order to get high scores.	3.77 (1.21)	3.52 (1.19)	-0.26	2.866	.005
6	I feel this test measures my English abilities correctly.	3.31 (1.13)	3.24 (1.01)	-0.07	0.949	.344
7	I want to take this test again if I have a chance.	2.91 (1.35)	2.99 (1.23)	0.07	-0.836	.404
8	I will participate more actively in speaking activities in an English class if I have to take this test again.	3.56 (1.18)	3.30 (1.04)	-0.26	3.070	.002
9	I tried to pronounce words more clearly because this test is computer-based.	3.22 (1.17)	3.45 (1.08)	0.23	-2.703	.007
10	I tried to pay attention to grammar when speaking because this test is computer-based.	2.67 (1.09)	2.88 (0.99)	0.21	-2.792	.006
11	I tried to speak more naturally because this test is computer-based.	2.91 (1.12)	3.09 (0.94)	0.18	-2.208	.028
12	Generally speaking, computer-based tests measure English abilities more correctly than person-based tests.	2.42 (1.03)	2.59 (1.03)	0.17	-2.300	.220
13	Generally speaking, I prefer computer-based tests to person-based tests when it comes to measuring my English abilities.	2.46 (1.07)	2.81 (1.24)	0.35	-4.120	< .001
14	How did you feel about taking this test? Please write freely.	-	3.24 (1.21)	-	-	-

Note. Adapted from "Daigaku eigo karikyuramu ni okeru gaibu supikingu tesuto no donju to shitsumonshi ni yoru datosei kensho," by Y. Shimizu and K. Owada, 2019, *Studies in Language Science Working Papers*, *9*, 1-14.

JACET Selected Papers Vol. 7 (2020), 160-173

A Four-Strand Approach to a College Listening-Speaking Course

Nozomu Sonda

Yamaguchi University

Abstract

It has been criticized that English education in Japanese schools has been heavily loaded with language-focused learning. As a result, many college students are not even ableto carry out basic conversation in English. Nation (2012) suggested a four-strand approach to language learning to balance our different types of learning and maximize their effects. These four strands are meaning-focused input, meaning-focused output, fluency development, and language-focused learning. According to Nation, each strand should have approximately one fourth of learning time allocated. The present study applied this approach to university English conversation classes in order to examines its effectiveness. As a result, it has been found that Nation's four-strand approach is quite effective. First, many students improved their fluency significantly and also gained more confidence in speaking English. Secondly, it was found that fluency and accuracy related activities can often go hand in hand; in fact, these two sometimes intertwined and mutually help each other. Thirdly, meaning-focused activities highly motivated the leaners and made learning much more fun. Fourthly, the four-strand approach allowed the instructor to see what was lacking or what was excessive in the teaching plans. The four-strand approach is highly recommended for college English classes.

Keywords: four-strand approach, fluency, accuracy, meaning

Introduction

It has been criticized that English education in Japanese schools has been heavily

loaded with language-focused learning such as grammatical drills, rote memorization of vocabulary, and intensive reading (Cripps, 2016). This is more so for that in junior and senior high schools because of the need to prepare for entrance exams. As a result, when they begin their college education, despite previous six years of English study, many students are not even able to carry out basic conversation in English. This is a serious educational problem of a national scale. Moreover, by that time, most of them will have developed inferior complex or aversion towards English. To be sure, the language-focused learning itself is a necessary component of English education; however, other learning activities should also be incorporated in order to achieve a well-balanced coherent curriculum. In relation to this, Paul Nation (2012, 2013a, 2013b, 2014, 2015; also see Nation & Yamamoto, 2012) suggested a four-strand approach to language learning; these four strands are meaning-focused input, meaning-focused output, fluency development, and language-focused learning. According to Nation, each strand should have approximately one fourth of learning time allocated. The present study applies this approach to university English conversation classes, examines the results, and makes recommendations for the betterment of English education in Japan.

Literature Review

Four-Strand Approach

Nation (ibid.) advocates the four-strand approach to language learning for a well-balanced, comprehensive language educational curriculum. The first strand is meaning-focused input. This involves listening and reading with a focus on the message. To enable the learners to focus on the meaning of the contents, the material needs to be easy for them. For example, 98% of the vocabulary in the reading material should be familiar to the reader. Activities can include listening to stories, paired conversation, extensive reading using graded readers, and extensive listening. The second strand is meaning-focused output, which involves

speaking and writing with the main attention to communication. The learners should speak or write about things they already know and very well. Relevant activities are self-introduction, paired conversations on casual topics, and presentations on topics of learners' interest. Fluency development is the third strand with a main focus on making use of what is already known to learners. For example, free conversation, speed reading/writing, and extensive listening/reading. Finally, language-focused learning consciously deals with the form of language such as grammar, vocabulary, discourse and pronunciation.

Research Questions

The research questions to be addressed in this study are as follows: (1) Does the four-strand approach to the college listening-speaking course effectively promote students learning? and (2) Is the time allocation for each strand appropriate?

Method

Participants and Classes

The participants in this study were those students who were enrolled in the present author's English conversation classes (upper level) at a national university in Japan. There were 7 classes of approximately 20 students each (total of 138 students), majoring in medicine, science, engineering, humanities, and agriculture. Their TOEIC scores are shown in a histogram on Figure 1 below. The average score is 551 with a standard deviation of 94.61. The test was taken at the end of the semester as part of the university requirements. The class was 90 minutes in duration per lesson and offered once a week for 15 weeks excluding a final exam week.

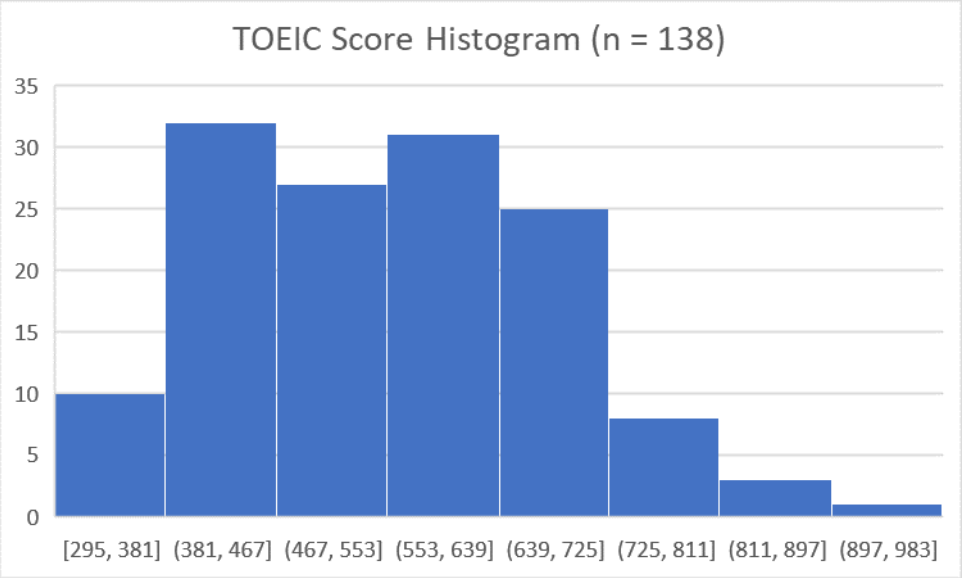


Figure 1. Histogram of TOEIC scores.

Lessons

The class was 90 minutes in duration per lesson and offered once a week for 15 weeks excluding a final exam week. In reality, however, due to the university's unified syllabus, the first week was spent on a test-taking skill lesson for the TOEIC test. In addition, on recommendation and also of the present author's own initiative, another lesson at the end of the semester was devoted on another test-taking skill lesson for both TOEIC and final exam (speaking test). Thus, there were actually 13 lessons which had the following pattern.

A typical lesson consisted of 3 sets of activities: self-introduction and warm-up conversation with partners (in pairs); useful words and expressions; back-channeling practice or presentations on social issues. At the end of the course, there was a self-evaluation/analysis sheet that the students filled out. The final exams consisted of two parts: one a test on handy expressions and back channeling and the other on monologue presentation on a chosen social issue.

Assessment Tools

The tools to assess the effectiveness of the four-strand approach are observation of the students' learning activities by the present author and analysis of student comments.

Results and Discussions

The first part of the lesson was always self-introduction, which consisted of 6 parts (See Appendix 1). Each part was used for two weeks, so over 12 weeks all 6 parts were covered. About 7 to 8 minutes were spent on this activity as a warm-up. Students were required to write out their self-introduction based on the sheets before the class. During this exercise, the present instructor did not intervene or interfere with the students some questions. This exercise was done to reinforce fluency and meaning (both input and output). This part was meant to be a light warm-up activity; however, the students enjoyed it more than expected. This is because students rarely got to talk to each other in class or hardly knew each other. According to them, most classes were lecture-style and there was little interaction in class. As a result, this part served as an excellent ice-breaker and an effective learning activity itself—learning about each other through English. It was therefore a very meaningful exercise and very useful in terms of meaning-focused input and output. It also facilitated fluency training because some of the items were recycled in every sheet or every now and then such as one's major, favorites, and unfavorites.

The self-introduction sheets progressed to be more challenging towards Parts 4,5 and 6, where questions were about the future of Japan and the world. Some students commented that these items were no longer part of -self-introduction but rather heavy discussion topics. However, the present author intentionally included these questions because they are very likely to be taken up as conversation topics even in initial encounters in the English speaking world.

The second part of learning some useful words and expressions for daily conversations.

This part was started with a review of the previous week's learning. The method used was an old-fashioned one; the instructor gave each students a few items to orally translate from Japanese to English within a second or two. Then a new set of expressions (about 20) were introduced with a few explanations, and the students worked in pairs to memorize these expressions. This part took about 15 to 20 minutes. The objective of this part was mostly fluency because the vocabulary and grammar of expressions used were quite simple and the main task of the students was to get used to them by hearing and speaking them repeatedly. There were a total of 10 lessons. The only challenge was many of the words and expressions were idiomatic and it would be difficult for students to come up with such expressions through mere translation.

The third part had two patterns, and each was used every other week. One pattern was to work on back channeling expressions (See Appendix 2). This exercise was adopted because the back channeling skills are very useful in terms of carrying and maintaining conversations but unfortunately not taught sufficiently in junior or senior high school. The first lesson was about making the most basic responses such as "Uh hum," "Really?," "Indeed!" and the like. The second and third lessons were about making a grammatical response based on the other person's speech. For example, if someone says, "I am studying science," then a correct response is "Yes, you are, aren't you?" This requires first a close attention to the original sentence structure, determining whether it is an affirmative or negative sentence. Secondly, the listener needs to determine the correct form of the pronoun; in this case, "I" becomes "you". If it is "my father," then it should be "he." If it is "my father, mother, and I," the correct pronoun is "you." Thirdly, the main verb of the original sentence must be identified with its correct tense, and then the correct form of the verb must be produced. For example, "I am studying" will be "you are" and "My father went" will be "he did" while "My mother has eaten" will be "she has." Finally, a tag question must be added based on the response sentence; if the response is affirmative, then the tag question must be negative while, if the

response is negative, the tag question must be affirmative. Thus, this exercise requires listening and grammatical skills. The fourth lesson is basically repeating what the other person says by either verbatim or paraphrasing or in a summarized form. This part is therefore both fluency and form-focused training.

The other pattern is presentations on social issues. Topics from contemporary social issues were taken, and the students were first required to prepare a 100- to 150-word essay on the selected topic, expressing their views and opinions about it. When they came to the class, they would be paired with a partner, working together to create a 2-minute conversation. They were given about 10 to 15 minutes to prepare it. They then presented their conversations. they repeated this pattern for 5 times and the last 3 times, the students made a 1-minute speech as a monologue based on the essay they prepared. The main focus of this exercise was meaning-based input and output as well as fluency training.

The lesson always ended with about 5 minutes for the students to write and submit their reflection on the day's activities. They wrote about what they learned, what was challenging for them, and what they hoped to do next as well as they asked questions and made suggestions for improvement.

Student Comments

Comments left by the students expressed their reflections on the learning activity, strong points and issues or challenges f the course. One wrote, “I paid little attention to accuracy throughout the semester. I aimed to understand what my partner said and respond to it at the same speed as I would in Japanese so that my message would be conveyed effectively.” Another commented, “I was very interested in the course work in that there were so many things that I had never put emphasis on such as handy expressions and back channeling. I think I overcame my lack of confidence in my speaking ability to a certain

degree." A third comment was, "I still have not gotten out of the study approach to English through mainly reading and writing, which I had been used to up through high school. So I want to put more focus on oral communication. For that purpose, I need to have more practical learning such as back channeling." Finally, there was this comment, "As I gave a short speech every week, I could really feel progress in speaking English more fluently. But I could never know whether my classmates really understood my speech or whether my contents and grammar were appropriate."

Implications

First, a four-strand approach allows you to see what is lacking or what is excessive in your teaching plans. In the present case, most of the students received instruction which emphasized accuracy up through their high school years, and therefore, fluency training and meaning-focused activities needed more attention. Based on this, an appropriate amount of time can be allocated to each strand. In the present case, when social issues were taken up, 30 minutes were allocated to fluency, 40 minutes to meaning, and 20 minutes to accuracy. When back channeling practice was pursued, 40 minutes were allocated to fluency, 30 minutes to meaning, and 30 minutes to accuracy.

Secondly, four 'strands' are not four separate lines. Rather, they are intertwined and may overlap with each other to a certain degree. For example, self-introduction was both fluency and meaning-based activities at the same time. Likewise, back channeling practice had both form-focused learning and fluency training as aims. Moreover, meaning-focused input and output activities often go together such as in self-introduction and social issues presentations. Thus, time allocation for each strand needs some flexibility and creativity in terms of calculation.

Through junior and senior high school English education, there is so much emphasis on accuracy but much less on fluency. The result is that most students are very shy or even afraid

of speaking English, let alone in front of class. However, even the output they make is not necessarily so great in terms of accuracy. Therefore, fluency training is much needed. In fact, fluency and accuracy are not opposing components of lessons; they rather go together hand in hand, reinforcing each other. Back channeling practice is a perfect example. As explained above, this practice requires close attention to the form of input. Grammatical rules involved are not too complicated; in fact, they are at the first- or second-year junior high school level. However, a great deal of training is required because there is so much information to be processed in such a short span of time. Through this practice, one's listening skills are improved, grammatical knowledge is strengthened, information processing speed is enhanced, and finally, output speed is accelerated. Thus, once you learn about the form, you should practice it to reinforce fluency.

There is no doubt that fluency is a severely lacking component in general English education in Japan. However, mere drill exercises to put knowledge into practice are not enough; in fact, they can be quite dull activities. Here the role of meaning-focused activity comes to be played. When there is meaning, learning is truly enhanced and becomes fun. As explained before, students are tired of sitting and listening to lectures all day; they need to talk to each other and learn about each other. They also want to discuss things of their interest, express their opinions and let out their emotions. For these reasons, self-introduction and social issues discussion can be very effective and productive.

Limitations

Although the students' speeches were evaluated every week either through back channeling practices or social issues presentations, there was only one rater, the present author. This may have affected the reliability of evaluation.

Students also commented that they felt their fluency improved significantly while

accuracy remained the same. This needs a closer analysis. Right now, fluency is catching up with accuracy, so we need more time. But the course is just one time offer. Although there isanother conversation course offered in the fall semester, instructors change, so there is no continuity of the lessons by the same instructor

Based on the student survey, students on the whole would like to have 50% on conversation and another 50% on social issues discussion although there was a clear tendency that the lower students preferred to have more time on daily conversation. But those who preferred to have more time on social issues is definitely a small minority. Thus, there is a greater need for daily conversation skills, but social issues do enhance the level of meaningfulness. For that matter, keeping the latter element can be justified.

After all, lack of time is an issue, which can truly be resolved at the level of the school's administrative. Having greater number of class levels, continuity of classes over the year, and possibly having two years of English classes as required—these might help. However, this is beyond the scope of this study and must yield to another occasion.

Conclusion

Nation's four-strand approach is quite effective. Although minor adjustment is always needed, all in all, the approximate one fourth for each strand seems quite valid. It is a highly recommended approach for college English classes. The present author is also writing an article on the same approach as applied to college reading-writing classes. However, much greater benefits may be accrued from application to junior and senior high school English classes, which can be an excellent research topic for the future.

References

Cripps, A. C. (2016). English Language Education in Japan: Problems and Solutions. In A. C.

Cripps (Ed.) *Perspectives on English Language Education in Japan*. Charleston, US:CreateSpace.Frank Smith, 1975, p. 162

Nation, P. (2012). What does every ESOL teacher need to know? Closing plenary address at the 2012 CLESOL conference in Palmerston North. TESOLANZ Journal, 20, 1-<http://www.tesolanz.org.nz/Site/Publications/TESOLANZ_Journal/>

Nation, I.S.P. and Yamamoto, A. (2012). Applying the four strands. *International Journal of Innovation in English Language Teaching and Research* 1(2), 167-181.

Nation, I.S.P. (2013a). *What should every EFL teacher know?* Seoul: Compass Publishing. ISBN 978-1-59966-266-4 <http://www.compasspub.com/english/book/book_view.asp?h_seq=1720>

Nation, I.S.P. (2013b). *What should every ESL teacher know*? Seoul: Compass Publishing.

Nation, P. (2014). *What you need to know to learn a foreign language*. <http://www.victoria.ac.nz/lals/about/staff/paul-nation>

Appendix 1: Self-Introduction (sample)

1. Hi, what's your name? （お名前は？）
2. My name is X. (I'm X.)　Call me Y. （わたしの名前は X です。 Y と呼んでください。）
3. Where are you from? （ご出身は？
4. I'm from Z Prefecture/City/Town/Village.　（Z 県・市・町・村出身です。）
5. What is your major? （専攻学科は何ですか？）
6. I major in XXX.　（XXX 専攻です。）
7. If you had plenty of time and money, where in Japan would you like to travel? Why? （時間とお金がたっぷりあったら、国内でどこへ旅行したいですか？なぜ？）
8. I want to go to XXX because…（XXX へ行きたいです。なぜならば．．．）
9. Where do you want to travel overseas? Why? （海外旅行はどこへ行きたいですか？なぜ？）
10. I want to go to XXX because…（XXX へ行きたいです。なぜならば．．．）
11. If you are able to study abroad, where do you want to go and what do you want to study? Why? （留学できるとしたら、どの国へ行って、何を勉強したいですか？なぜ？）
12. I want to go to XXX to study YYY because …（XXX へ行って、YYY を勉強したいです。なぜならば．．．）。
13. If you can live abroad and work, where would you go? What would you do? Why? （外国に住んで働けるとしたら、どこへ行きますか？何の仕事をしますか？なぜ？）
14. I would go to XXX and do YYY because …（XXX へ行って、YYY をしたいです。なぜならば．．．）。
15. If you could speak and write English fluently, what would you do with English? （もし英語［もっと］流暢に話したり書いたりできたら、英語を使って何をしたいですか？）
16. I would …（．．．．したいです）

Appendix 2: Back Channeling (sample)

すべてボックス内の相槌４のパターンで答える。ただし、1－16 のステートメントに対しては、肯定文は肯定で、否定文は否定で答える。28-34 は、相槌を打つ側が判断して肯定にするか否定にするか決めて答える。確認事項＝主語（必ず代名詞に置き換える）、動詞の種類（一般動詞・be 動詞・助動詞）、時制、イントネーション。イントネーションは、疑問がほとんど、あるいはまったくなければ下げる、疑問がいくらか、あるいはかなりあれば上げる。

1. Hi, I am (　　).
2. I come from (　　).
3. I went to (　　) last weekend.
4. I was very happy.
5. I have been to (　　).
6. But I haven't been to (　　).
7. I would like to visit (　　) someday.
8. I will become a/an (　　) in the future.
9. I should study (　　) very hard, though.
10. I mustn't goof around. （だらだらしていてはいけない）
11. I should go now.
12. My mother is a (　　).
13. When young, my mother was a/an (　　).
14. My mother did a lot of (　　)ing then.
15. Now my mother wants to become a/an (　　).
16. My father is a/an (　　).

相槌１：最もカジュアルな相槌

Uh huh. Uh hum. Yep. Right. Exactly.

That's right. You're right. Indeed. Wow.

Really? Super. Great. No kidding.

Unbelievable. Incredible. Awesome.

相槌２：上級相槌(疑問形語順)
Do you? Is it? Has she? Did he? 等々

相槌３：上級相槌 A(通常語順)
You do? It is? She has? He did? 等 々

相槌４:上級の上級（2 と 3 の合体編、Yes/No で始める）

Yes, you do, don't you? No, he didn't,

17. When young, my father was a/an ( ).

18. My father did a lot of ( )ing then.

19. Now my father wants to become a/an ( ).

20. My mother and my father are a happy couple.

21. They like to go out together.

22. They have been to ( ) before.

23. When young, they did a lot of things together.

24. But they've never been to ( ).

25. They would like to visit ( ), too.

26. Last year, my parents and I went to ( ).

27. My parents and I are going to ( ) this winter.

****

28. You speak very good English.

29. When in high school, you weren't good at English.

30. Now, you are studying Chinese, too.

31. You should become a teacher.

32. But you must study much harder.

33. Then you will be very successful in the future.

The Effects of Informing the Quality of Students' Previous Peer Assessment

Sumie Matsuno

Aichi Sangyo University College

Abstract

In recent years, peer assessment has become of particular interest in educational assessment with the increasing emphasis on active learning. Although a considerable number of studies have shown that peer assessment is a good pedagogical method for teaching language, there is little agreement as to if it can be used to supplement teacher assessment as a formal assessment; hence, this study explores one way to lead toward more sophisticated peer assessment. In two English classes, university students gave two presentations: the first presentation (FP) and the second presentation (SP). In the experimental group (EG), after FP, peer assessments were analyzed using the FACETS computer program, and each student received comments from their teacher about the quality of their peer assessment. Meanwhile, in the control group (CG), students did not obtain any comments between FP and SP. This study concerned how those comments affected peer assessments of SP. The results indicated that, in EG, unexpected responses dramatically decreased and the wider variety of score categories were utilized in SP. Moreover, the bias interactions between raters and items of the EG also significantly decreased. Overall, the study indicated the effectiveness of indicating how students had undertaken their previous peer assessment.

***Keywords*:** FACETS, peer assessment, presentation, university

Introduction

Peer assessment has received significant attention in studies on English as a second language (ESL) and English as a foreign language (EFL) for many years. Numerous attempts have been made by scholars to show how peer assessment has educational benefits for students, such as improving their own performance, enhancing autonomy, and assessing their own language skills (Almahasneh & Abdul-Hamid, 2019; Bowman-Perrott, Mahadevan, and Etchells, 2016; Cheng & Warren, 2005; Colby, 2011; Shin, 2013; Jingjing & Fengying, 2015; Iraji, Enayat, & Momeni, 2016). On the other hand, there is little agreement as to whether peer assessment works as a formal assessment (Azarnoosh, 2013; Birjandi & Bolghari, 2015; Cheng and Warren, 1999; Farrokhi, Esfandiari, and Schaefer, 2012; Güler, 2017; Jafarpur, 1991; Jones and Alcock , 2014; Matsuno, 2009; Patri, 2002; Panadero & Alqassab, 2019; Saito, 2008; Saito & Fujita, 2004; Topping, 2009; Vanderhoven et al., 2015). Hence, this paper is intended as an investigation of whether the quality of peer assessment can be improved in a Japanese university classroom after informing students how each of them assessed their peers in the previous assessment and advising them how they can improve their peer assessment in the next presentation.

Literature Review

An abundance of studies have suggested that peer assessment is a good pedagogical method for teaching language. Cheng and Warren (2005) indicated that "…if language learners could be trained to confidently and reliably assess the language proficiency of their peers, they would also be able to confidently evaluate their own language skills, a valuable precondition for improving them" (p.110). According to Colby (2011), many learners felt that they had learned with their peers, and that peer assessment contributed to learning. Shin (2013) noted that students could improve their specialized knowledge through peer assessing and could also significantly enhance their learning motivation and interest. Reviewing some research results, Afitska (2014) mentioned that peer assessment may

enhance learner autonomy by allowing said learners to take control and assume ownership of their learning while also recognizing that they have responsibility for their own learning. In the study conducted by Jingjing and Fengying (2015), students found peer assessment helpful in improving their performance, and especially they liked receiving the grading sheet from their peers. Ahmad and Majid (2015) showed that peer assessment was a supplementary tool for students' engagement and empowerment. Reviewing 17 studies, Bowman-Perrott, Mahadevan and Etchells (2016) noticed that peer assessment encouraged EFL learners of various levels of English proficiency to improve their English. Moreover, Casallas and Samir (2016) found that peer assessment enhanced personal and collaborative learning and promoted learner reflection and agency. Iraji, Enayat and Momeni (2016) revealed that students showed a positive attitude towards peer assessment which facilitated them learning from the mistakes of their peers. Almahasneh and Abdul-Hamid (2019) proved that students who had received peer assessment training wrote a better quality of essay than students who only took conventional essay writing training.

On the other hand, there has been debate regarding whether peer assessment is an effective measurement tool. Cheng and Warren (1999) revealed that the peer assessments are not reliable enough to be used to supplement teacher assessment; however, they found some evidence that practical experience of assessing a particular task type can improve students' assessment skills when they assess a similar task. Farrokhi, Esfandiari and Schaefer (2012) claimed that the patterns of severity and leniency were opposite between teacher raters and peer raters when assessing the highest and lowest ability students. In contrast, some studies showed that peer assessment was correlated with teacher assessment (Jafarpur, 1991; Patri, 2002; Saito & Fujita, 2004). Topping (2009) mentioned, "a peer assessor with less skill at assessment but more time in which to do it could produce an equally reliable and valid assessment" (p. 25). Saito (2008) showed the high correlation between teachers' ratings and peers' ratings after training peers to assess. Further, Matsuno (2009)

found that peer assessment was internally consistent and showed few bias interactions. Azarnoosh (2013) discovered no significant difference between the learners' peer assessment and teacher assessment, and no friendship bias was found in peer assessment. Jones and Alcock (2014) found high validity and inter-rater reliability by asking students to compare pairs of scripts against one another and concluded that the students performed well as peer assessors. Birjandi and Bolghari (2015) found that peer- and self-rating could be valid and reliable assessment tools because the high correlations were obtained between teacher-, self-, and peer-assessment. Vanderhoven et al. (2015) discovered that grades achieved by students in the non-anonymous condition were highly correlated to the teachers' grades, while Güler (2017) uncovered a high correlation between them in the anonymous condition. Panadero and Alqassab (2019) reviewed the studies of peer assessment, focusing on anonymity and no-anonymity, and found mixed results regarding whether peer assessment was effective.

As seen in the above, while peer assessment is now widely accepted as a good pedagogical method, there is still room for a considerable measure of disagreement about peer assessment as an effective measurement tool. However, if peer assessment becomes more reliable, it can be used to supplement teacher assessment. This study is conducted to establish whether the quality of peer assessment can be improved in a Japanese university classroom after informing students how each of them assessed their peers in the previous assessment and advising them how they can improve their peer assessment in the next presentation. Multifaceted Rasch analysis is conducted since, utilizing this analysis, rater factors can be investigated and feedback to raters on their rating performance can easily be provided. Six research question (RQ) are explored.

RQ1: In the control group (CG) and experimental group (EG), to what degree could presenters' abilities, raters' severities, and items' difficulties be used to be measured?

RQ2: In each group, to what degree did raters' severities differ between the first presentation (FP) and the second presentation (SP)?

RQ3: In each group, how did the number of misfitting raters, presenters, and items change between FP and SP?

RQ4: In each group, how did category statistics change between FP and SP?

RQ5: In each group, how many unexpected responses were detected at FP and SP?

RQ6: In each group, how did the rater biases change between FP and SP?

Methods

Procedure

In two English classes, university students took a required English class and gave two presentations, namely the first presentation (FP) and second one (SP). Their average of TOEFL scores was around 490. Each presentation was approximately 3 minutes long and was in English. There was a 3-week interval between FP and SP. During those presentations, the students assessed their peers' presentations based on six items: posture & eye contact, gestures & voice inflection, visuals, fluency & pronunciation, and content & organization, and overall. In both FP and SP, each item was scored on a 5-point Likert scale. The rater assigned a score of 1, 2, 3, 4, or 5 representing a presentation that ranged from "very poor", "poor", and "OK", to "good", and "excellent".

In each class, 19 students participated in the present study. They were both presenters (P1-P19) and raters (R1-R19). In both groups, before the presentation, the same teacher explained what these six items were and how to rate the presentations thoroughly for approximately three successive classes (each class lasted 90 minutes). They watched the model presentation and were given an explanation on how to give a good presentation and evaluate peers' presentations.

In the experimental group (EG), where the raters were from R1 to R19 and the presenters were

from P1 to P19, after FP, their peer assessments were analyzed using the FACETS computer program, version 3.80.4 (Linacre, 2018), and each student received some comments from his/her teacher regarding the quality of his/her own peer assessment based on the results of FACETS analysis. Those comments included his/her severity, his/her unexpected responses, his/her use rate of categories, and his/her biases. On the other hand, in the control group (CG), where the raters were from R1 to R19 and the presenters were from P1 to P19, they did not obtain any comments between FP and SP. The present study concerns how those comments affected their peer assessment of SP.

Analysis

Multifaceted Rasch analysis was conducted using the FACETS computer program, version 3.80.4 (Linacre, 2018). In this study. rater severity, presenter ability, and item difficulty were the facets. The output of the FACETS analysis reported summary statistics of raters, presenters and items to check whether they were in the acceptable range to be measured. It also displayed a FACETS map, which showed visual information about differences that existed among those facets. Presenter ability logit measures were estimated simultaneously with the rater severity logit estimates and item difficulty logit estimates on the same linear measurement scale (logit scale), which meant they were easy to compare. It also reported an ability measure and fit statistic for each presenter, a severity measure and fit statistic for each rater, and a difficulty estimate and fit statistic for each item. In addition, it showed unexpected responses, which might cause presenters, raters, or items to be misfitting. Moreover, bias analyses were carried out in order to detect any raters who were rating particular items too severely or too leniently.

The FACETS analysis was run four times: FP of CG, FP of EG, SP of CG, and SP of EG. Based on the results of FP of EG, each student in EG received some comments from his/her teacher. Hence, it was necessary to run the FACETS for FPs and for SPs separately. Moreover, running four

FACETS made us easier to interpret the results of how CG and EG changed between FP and SP. In this study, only peer assessment was analyzed by FACETS.

Results

Summary statistics

The first research question was "In the control group (CG) and experimental group (EG), to what degree could presenters' abilities, raters' severities, and items' difficulties be used to be measured?" To answer this question, the following were the summary statistics of the first and the second presentations of CG in Table 1 and those of EG in Table 2. Each table illustrates the first presentation/the second presentation.

Table 1

Summary Statistics of the First Presentation and the Second Presentation (the first/the second) in the Control Group (CG)

	Presenters	Raters	Items
Logic-Scale			
Infit MSE			
M	1.00 / 1.00	1.03 /1.00	1.01 / 1.00
SD	.28 / .19	.43 / .35	.29 / .19
Std.Infit MSE			
M	- .10 / .00	- .10 / - .20	- .10 / - .10
SD	1.90/ .12	2.90 / 2.60	3.40 / 2.50
Outfit MSE			
M	.98 / 1.00	.96 / .99	.98 / 1.00

SD	.31 / .19	.34 / .34 /	.28 / .19
Std. Outfit MSE			
M	- .20/ .00	- .50 / - .30	- .40 /- .20
SD	2.00/ 1.20	2.40 / 2.50	2.90 / 2.50
RS	.98 / .98	.98 / .98	.94 / .93
Chi-square	778.80* /760.30*	1020.70* /787.50*	83.30* / 67.30*
Degree of Freedom	18 / 18	18 / 18	5/ 5

Note. RS= reliability of separation.

Table 2

Summary Statistics of the First Presentation and the Second Presentation (the first/the second) in the Experimental Group (EG)

	Presenters	Raters	Items
Logic-Scale			
Infit MSE			
M	.99/ .99	1.05/ 1.03	1.00 / 1.00
SD	.27 / 0.23	.27 / .34	.27 / .21
Std.Infit MSE			
M	- .20/ - .10	.10 / - .10	- .30 / - .20
SD	1.80/ 1.60	1.80 / 2.40	3.30/ 3.10
Outfit MSE			
M	.99 / .99	.99 / 1.02	.99 / .99
SD	.36 / .23	.22 / .31	.29 / .21
Std. Outfit MSE			

M	.00/ - .10	- .20 / - .20	- .20 / - .30
SD	.00 / 1.60	1.40 / 2.30	2.50 / 2.90
RS	.97 / .99.99 /	.99 .95 / .95	
Chi-square	564.40* /567.80*	1640.30* / 1260.40*	95.80* / 101.10*
Degree of Freedom	18 / 18	18 / 18	5 / 5

Note. RS= reliability of separation.

Linacre (2012) mentioned that the values from .5 to 1.5 (the logit scale) of infit and outfit mean-square statistics are "productive for measurement" (Linacre, 2012, p.15). In these tables of CG and EG, all of the presenters, raters, and items seemed to be acceptable; that is, the data could be utilized to perform measurements. In addition, the acceptable range of standardized fit statistics is between -1.9 and 1.9, which indicates "data have reasonable predictability" (Linacre, 2002, p.878). Again, all of the standardized infit and outfit mean square values were within the acceptable range of between -1.9 and 1.9, which also indicated that the data could be used to be measured.

High reliabilities of separation statistics for presenters, raters and items suggested that there were reliable differences in the judged locations of each presenter's ability, each rater's severity, and each item's difficulty respectively, on the logit scale. Varying raters' severities was not ideal for raters; however, this is often the case in real classroom settings.

A chi-square statistic determines whether the element within a facet can be exchangeable. The overall differences between elements within the presenter, rater, and item facets were significant, based on the chi-square statistics ($p < .05$). The noteworthy point was that in EG, the presenter facet had a slightly larger chi-square value on the second presentation, meaning that the students showed more different abilities than the first presentation, while in CG the presenters' chi-square value became lower on the SP than FP. That is, the students in EG differentiated their peers' writing ability

better than those in FP.

The FACETS Maps

The second research question was "In each group, to what degree did raters' severities differ between FP and SP?" To answer this question, the FACETS maps (Figure 1 and Figure 2) were presented. In all of the FACETS maps, the first column was the logit scale, to which all the scores were converted. The next three columns displayed the logit-scale locations for the three facets: raters, presenters, and items. In order to interpret the logit-scale locations of the three facets, raters and items were centered at zero (mean set to zero), and only the average location of the presenter facet was allowed to vary. The second column displayed the raters' severity locations (n=19). The raters located higher on the logit scale were more severe; that is, they assigned lower ratings more often. The third column showed presenters' ability locations; those located higher were more able presenters. The fourth column showed the locations of the items; the items located higher on the logit scale were associated with more severe ratings, and the items located lower on the logit scale were associated with less severe ratings. The fifth column showed the scores (categories) that the raters actually assigned.

Figure 1 shows the FACETS map of FP of CG. R14 was the most severe; however, even R14's severity could not differentiate the presenters' abilities well enough. Moreover, R14 could not segregate the abilities of P16, P9, P19, P17, P6, P13, and P11. On the other hand, R10, R11, R1, R13, R15, R12, R19, and R4 were too lenient raters. They gave all the presenters good scores, so their ratings were not reliable.

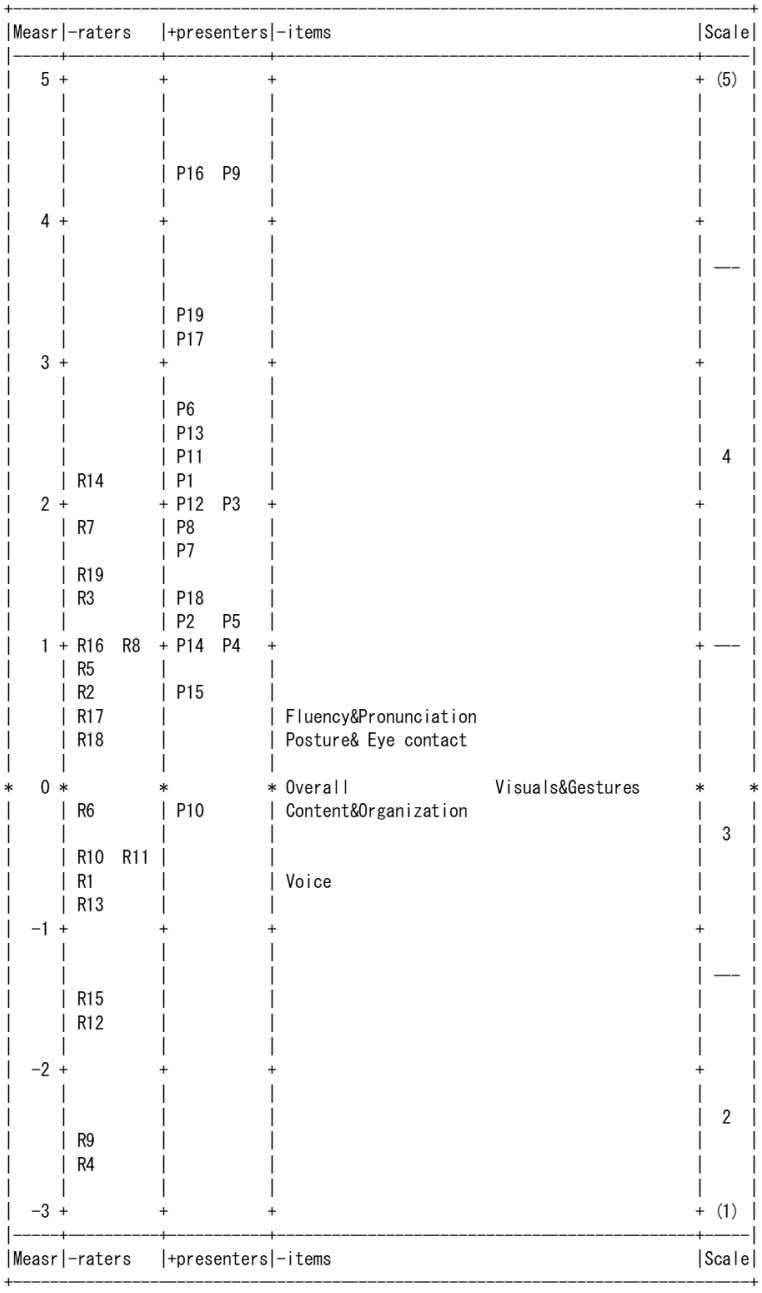


Figure 1. The FACET Map of FP of CG.

Figure 2 shows the FACETS map of SP of CG. R14 was the most severe rater again; he/she could not distinguish the abilities of P6 and P7 well. Since at FP, he/she could not distinguish seven presenters' abilities, it seemed that he/she improved his/her ratings. In this presentation, P3's ability was very low, and so all the raters except R4 could rate P3's ability correctly.

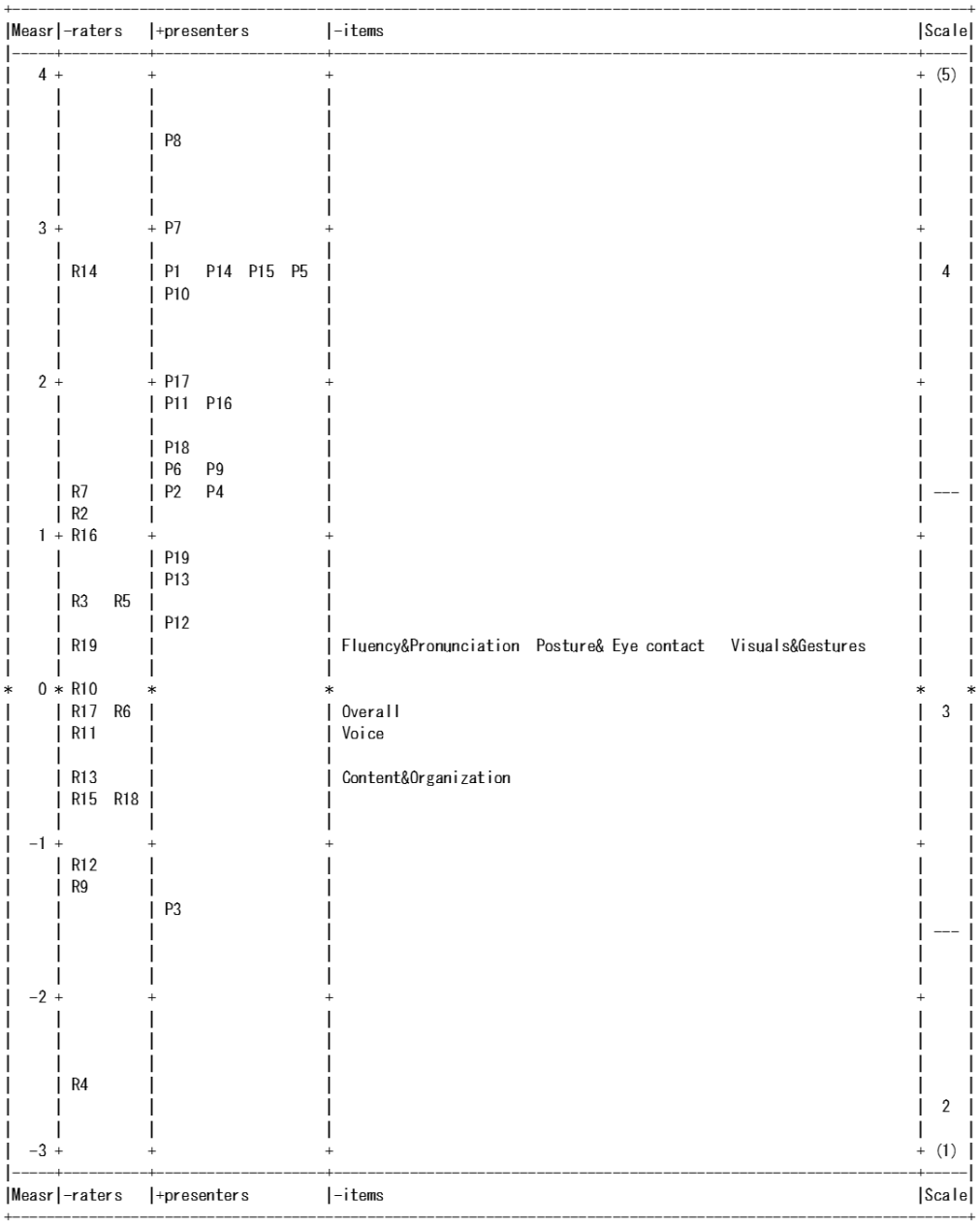


Figure 2. The FACET Map of SP of CG.

Figure 3 presents the FACET map of FP of EG. Although R7 was the most severe rater, he/she could not evaluate P16's ability well. On the other hand, R2, R6, R11, R3, R18, R9, R4, R15, R13, R8, and R16 were too lenient to differentiate all the presenters' abilities. Their ratings were not reliable.

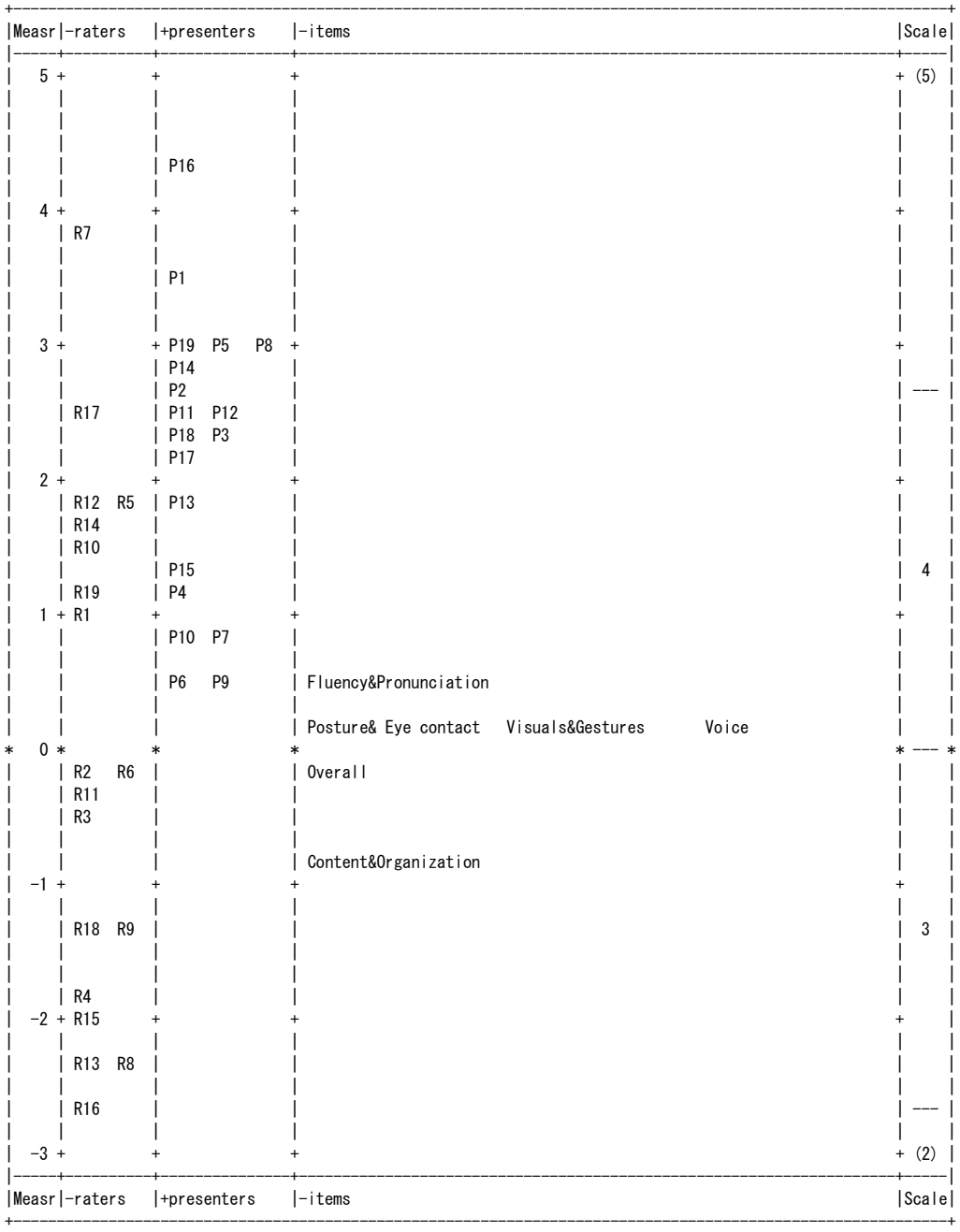


Figure 3. The FACET Map of the First Presentation of EG.

Figure 4 illustrates the FACETS map of SP of EG. R14 could evaluate all the presenters' abilities well; however, R18, R15, R2, R4, R17, R6, R8, and R13, who were too lenient, could not distinguish all the presenters' abilities.

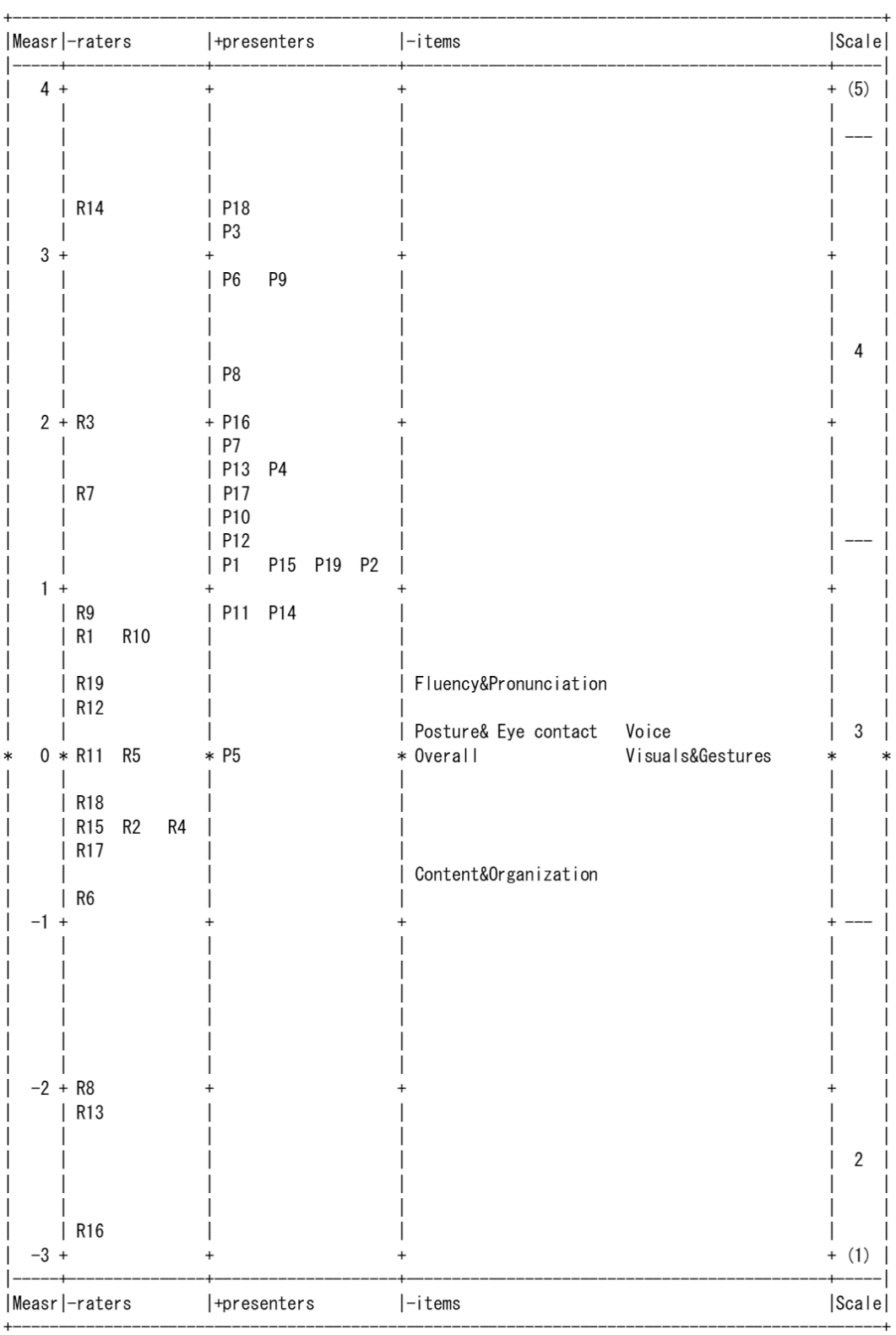


Figure 4. The FACET Map of the Second Presentation of EG.

The following is a summary of Figure 1 to Figure 4.

Table 3

Summary of FACETS maps

	FP of CG	SP of CG	FP of EG	SP of EG
The number of too high ability presenters	7	2	1	0
The number of too lenient raters	8	1	11	9

Too high ability presenters are those whose abilities could not be distinguished well by any raters; too lenient raters are those who could not distinguish the abilities of any presenters. We can notice that both CG and EG could improve the quality of peer assessments since those numbers decreased from FP to SP.

Misfitting Raters, Presenters, and Items

Research question 3 was "In each group, how did the number of misfitting raters, presenters, and items change between FP and SP?" Table 4 presents the answer to this question.

Table 4

The Number of Misfitting Raters, Presenters, and Items

	Control group (FP→SP)	Experimental Group (FP→SP)
The number of misfitting raters	4→1	1→2
The number of misfitting presenters	1→0	1→2
The number of misfitting items	1→0	1→0

While the number decreased in CG, the number of misfitting raters and presenters increased in EG. This seems to lead to the conclusion that CG assessed their peers better than EG at SP.

Tables 5, 6, 7 and 8 represent the details of misfitting raters.

Table 5

The Misfitting Raters in the Control Group (FP)

Rater	Severity	Error	Infit mean Square	Outfit mean Square
R4	-2.70	.24	1.84	1.28
R9	-2.45	.20	1.66	1.22
R15	-1.58	.18	1.61	1.36
R13	-.77	.16	.44	.45
R18	.42	.16	.45	.49
R5	.76	.15	1.63	1.60

Table 6

The Misfitting Raters in the Control Group (SP)

Rater	Severity	Error	Infit mean Square	Outfit mean Square
R4	-2.59	.19	1.59	1.44
R18	-.71	.20	0.37	.37
R5	.55	.16	1.76	1.78

Rater	Severity	Error	Infit mean Square	Outfit mean Square
R15	-2.06	.24	1.60	1.18

From the left, each column showed rater IDs, rater severity, error, infit mean square values, and outfit mean square values. On the first presentation, six raters were detected as misfit; on the other hand,

with regard to the second presentation, three raters were detected as misfit. All three misfitting raters were also misfit at the first presentation.

The following are the results of the experimental group.

Table 7

The Misfitting Raters in the Experimental Group (FP)

Rater	Severity	Error	Infit mean Square	Outfit mean Square
R15	-2.06	.24	1.60	1.18

Table 8

The Misfitting Raters in the Experimental Group (FP)

Rater	Severity	Error	Infit mean Square	Outfit mean Square
R13	-2.18	.27	1.78	1.57
R15	-.45	.15	1.52	1.52

In EG, after informing their quality of peer assessment at FP, the number of misfitting raters increased. Moreover, R15 was still misfitting at SP. From this result it seems that informing the previous peer assessment did not contribute to improving R15's peer assessment.

Category Statistics

Research question 5 was "In each group, how did category statistics change between FP and SP?" This question was answered by the category statistics (Table 9) presented in FACETS.

Table 9

The Category Statistics

Category	CG (FP→SP)	EG (FP→SP)
1	1% → 1%	0% → 1%
2	4% → 6%	4% → 11%
3	28% → 34%	17% → 34%
4	40% → 48%	36% → 36%
5	27% → 12%	44% → 19%

Using various categories usually leads to successfully differentiating learners' abilities, which is the main purpose of the assessment. In Table 9, in EG at FP, 44% of all the categories they utilized were category 5; however, this decreased from 44% to 19%, which was a remarkable change. In both CG and EG, their utilization of categories improved from FP to SP because they came to use a wider variety of categories. However, this tendency was more prominent in EG than CG.

Unexpected Responses

The fifth research question was "In each group, how did category statistics change between FP and SP?" To answer this question, it is important to look at Tables 10 to 13. In Table 10, for example, when R4 rated P7's item, he/she gave the category 2; however, considering R4's severity, P4's ability, and the difficulty of fluency, P4 should have received 4.6. Hence, this was an unexpected response. In FP, CG had 12 unexpected responses, with this number decreasing to 7 responses. With regard to EG, they had 27 unexpected responses, which dramatically decreased to 4. The decrease in number was more explicit in EG than CG.

Table 10

Unexpected responses of FP (CG)

The number of unexpected responses

Rater	Presenter	Item	Score	Expected
R4	P7	Fluency	2	4.6
R9	P2	Visuals	2	4.5
R9	P14	Visuals	2	4.4
R5	P14	Visuals	1	3.2
R6	P10	Visuals	1	3.2
R5	P18	Posture	1	3.1
R15	P7	Fluency	2	4.2
R4	P14	Voice	3	4.7
R7	P8	Posture	3.0	2.0
R11	P19	Content	3.0	4.6
R4	P15	Voice	3.0	4.6
R5	P15	Posture	1.0	2.9

Table 11

Unexpected responses of SP (CG)

The number of unexpected responses				
Rater	Presenter	Item	Score	Expected
R4	P3	Fluency	1	3.4
R14	P8	Posture	1	3.3
R14	P4	Visuals	1	3.0
R5	P19	Posture	5	3.1
P13	P3	Voice	1	2.9
R2	P18	Visuals	5	3.1

Table 12

Unexpected Responses of FP (EG)

The number of unexpected responses				
Rater	Presenter	Item	Score	Expected
R8	P8	Content	4	5.0
R16	P3	Fluency	3	4.9
R13	P7	Content	3	4.8
R8	P12	Content	4	5.0
R15	P17	Visuals	3	4.8
R8	P13	Visuals	3	4.8
R9	P19	Content	4	4.9
R9	P8	Content	4	4.9
R3	P13	Visuals	2	4.3
R15	P13	Visuals	3	4.7
R20	P11	Posture	2	4.3
R13	P7	Overall	3	4.7
R16	P3	Overall	4	4.9
R20	P17	Visuals	2	4.2
R2	P13	Visuals	2	4.2
R4	P1	Voice	4	4.9
R2	P6	Content	2	4.1
R8	P13	Content	4	4.9
R20	P13	Visuals	2	4.1

R2	P9	Content	2	4.0
R9	P2	Content	4	4.9
R15	P19	Visuals	4	4.9
R3	P7	Posture	2	4.0
R19	P1	Content	3	4.6
R3	P3	Fluency	2	3.9
R20	P15	Posture	2	3.9

Table 13

Unexpected Responses of SP (EG)

The number of unexpected responses				
Rater	Presenter	Item	Score	Expected
R16	P6	Content	4	5.0
R3	P16	Voice	5	2.8
R4	P8	Voice	2	4.1
R10	P1	Voice	1	3.0

Bias Analysis (Rater × Items)

The sixth research question was "In each group, how did the rater biases change between FP and SP?" This question was answered by looking at the information contained within Tables 14 to 17, which show the bias interactions between peer raters and items. Rater bias occurs when a particular rater assesses a specific item too severely or too leniently. In this table, bias interactions with a probability of less than .05 are presented, meaning the null hypothesis of "the same severity/leniency" was rejected at $p < .05$; in other words, those raters had statistically significant

bias.

In Table 14, for example, when R9 rated Visuals, the observed score was 76, but based on R9's overall severity and the difficulty of Visuals, we expected a summed score near 83.81. Hence, R9 gave Visuals 7.81 rating-points lower than expected. This bias size was -1.48 logits, with the model standard error of .40. A test of the hypothesis "this bias is due to measurement error" yielded a t = -3.70. The null hypothesis of "the same severity" was rejected at $p < .05$. To be more precise, this bias (R9 vs. Visuals) was statistically significant.

Table 14

Bias Report: FP of CG

Rater	Item	Observed	Expected	Average-observed-expected raw score	Bias	Error	*t*	*p*
R9	Visuals	76	83.81	- .43	-1.48	.40	-3.70	.001
R17	Voice	62	70.19	- .46	-1.19	.38	-3.15	.005
R4	Visuals	67	71.06	- .27	-1.14	.48	-2.38	.032
R8	Visuals	58	52.65	.36	.93	.42	2.20	.044
R6	Voice	81	75.16	.32	1.03	.45	2.30	.034
R2	Content	75	67.34	.43	1.15	.40	2.89	.010
R17	Visuals	75	66.52	.47	1.27	.40	3.20	.005
R9	Posture	88	81.82	.34	1.92	.77	2.51	.022

Table 15

Bias Report: SP of CG

Rater	Item	Observed	Expected	Average-observed-expected raw score	Bias	Error	*t*	*p*
R17	Voice	61	68.69	- .43	-1.25	.40	-3.13	.006
R19	Fluency	52	58.43	- .38	-1.07	.41	-2.65	.017
R2	Posture	50	56.89	- .38	-1.06	.39	-2.73	.014
R17	Posture	59	64.63	- .31	- .90	.40	-2.26	.037
R14	Posture	41	46.78	- .32	- .83	.38	-2.21	.041
R19	Posture	64	58.54	.32	.94	.42	2.24	.013
R2	Content	69	62.23	.38	1.11	.41	2.70	.002

Table 16

Bias Report: FP of EG

Rater	Item	Observed	Expected	Average-observed-expected raw score	Bias	Error	*t*	*p*
R9	Content	81	86.74	- .31	-1.54	.46	-3.37	.003
R8	Content	81	88.49	- .19	-1.53	.54	-2.81	.012
R18	Visuals	78	83.90	- .33	-1.23	.42	-2.92	.009
R4	Fluency	78	83.87	- .33	-1.22	.42	-2.91	.009
R1	Voice	60	66.61	- .39	-1.04	.39	-2.64	.017
R15	Fluency	80	84.50	- .25	-1.03	.44	-2.33	.032
R15	Visuals	82	85.68	- .20	- .98	.47	-2.07	.053
R17	Posture	53	59.22	- .35	- .92	.39	-2.35	.031

R3	Posture	74	78.84	- .27	- .83	.40	-2.07	.054
R19	Visuals	74	68.36	.31	.87	.40	2.17	.044
R10	Fluency	70	63.94	.34	.90	.39	2.30	.034
R11	Posture	83	78.10	.34	.90	.39	2.30	.056
R1	Posture	74	66.41	.45	1.31	.44	2.98	.008
R3	Content	89	83.77	.29	2.34	1.05	2.24	.039
R9	Voice	90	83.26	.37	3.61	1.74	2.07	.053

Table 17

Bias Report: SP of EG

Rater	Item	Observed	Expected	Average-observed-expected raw score	Bias	Error	*t*	*p*
R6	Visuals	63	71.89	- .49	-1.14	.36	-3.20	.052
R6	Content	67	73.50	- .38	- .97	.37	-2.59	.019
R15	Fluency	59	66.30	- .41	- .93	.36	-2.58	.019
R11	Voice	57	63.94	- .39	- .88	.36	-2.46	.025
R15	Visuals	77	69.19	.43	1.07	.39	2.76	.013
R11	Content	81	71.02	.55	1.47	.42	3.47	.002

In CG, the number of biases decreased from 8 to 7, while in EG this number decreased from 15 to 6. Conversely, in both CG and EG, biases decreased from FP to SP, and EG improved peer-assessments more than CG.

Discussion and Conclusion

The first research question was "In the control group (CG) and experimental group (EG), to what degree could presenters' abilities, raters' severities, and items' difficulties be used to be measured?" The summary statistics of Rash analyses show that the data could be used to be measured. The second research question was "In each group, to what degree did raters' severities differ between the first presentation (FP) and the second presentation (SP)?" Both CG and EG could improve the quality of peer assessments between FP and SP since the number of indistinguishable presentation abilities decreased. The third research question was "In each group, how did the number of misfitting raters, presenters, and items change between FP and SP?" In EG, after informing their quality of peer assessment at FP, the number of misfitting raters increased. Moreover, one misfitting rater at FT was still misfitting at SP, which indicated that informing the previous peer assessment did not lead to improving his/her peer assessment. The fourth research question was "In each group, how did category statistics change between FP and SP?" In both CG and EG, their utilization of categories improved between FP and SP because they came to use a wider variety of categories; however, this tendency was more prominent in EG than CG. This showed informing their quality of peer assessment might contribute to improve their assessment. In the fifth research question was "In each group, how many unexpected responses were detected at FP and SP?" In EG, the number of unexpected responses dramatically decreased from 27 to 4. This could also indicate informing the previous assessment could improve their quality of peer assessment. The final question was "In each group, how did the rater biases change between FP and SP?" The results showed that rater biases dramatically decreased in EG from 15 to 6, while those decreased in CG from 8 to 7.

In summary, though both CG and EG improved their peer assessments from FP to SP, it seemed that EG improved more than CG with regard to three points. First, unexpected responses decreased from FP to SP, which was more prominent in EG. Second, students used wider varieties of

categories in SP than in FP, which was more remarkable in EG. Third, rater biases dramatically decreased in EG. Those results showed that informing the quality of their previous peer assessment could lead to better peer assessment at the next presentation. Therefore, even a small amount of information about their peer assessment, such as informing what categories they frequently utilized, should be told to them to make their assessment more efficient.

Limitations

There were some limitations in this study. First, the number of participants was too small; only 38 joined this study. Second, the level of the university studied was upper-intermediate. The results of the present study may not be generalized to students of other levels. Third, both CG and EG improved FP to SP, and so without informing the quality of students' previous peer assessment, they might improve their quality of peer assessments. Though there are some limitations, it is hoped that this study will pave the way for improving students' peer assessment.

Acknowledgement

I would like to express my gratitude to the three anonymous reviews for their helpful comments to clarify and strengthen the final draft for this article. Needless to say, any errors that may remain are solely my responsibility.

References

Afitska, O. (2014). Use of formative assessment, self and peer-assessment in the classrooms: some insights from recent language testing and assessment research. *i-Manager's Journal on English Language Teaching,* 4(1), 29-39.

Ahmad, Y., & Majid, M. (2015). Portfolio assessment, peer assessment and writing skill

improvement. *Theory and Practice in Language Studies*, 5(12), 2504-2511.

Almahasneh, A.M.S., & Abdul- Hamid, S. (2019). The Effect of Using Peer Assessment Training on Writing Performance among Arab EFL High School Students in Malaysia. *Arab World English Journal,* 10 (1)105-115. Retrieved from https://dx.doi.org/10.24093/awej/vol10no1.10.

Azarnoosh, M. (2013). Peer assessment in an EFL context: attitudes and friendship bias. *Language Testing in Asia*, 3(11), 1-10.

Birjandi, P., & Bolghari, M.S. (2015). The relationship between the accuracy of self- and peer-assessment of Iranian intermediate EFL learners and their learning styles. *Theory and Practice in Language Studies*, 5(5), 996-1006.

Bowman-Perrott, L., Mahadevan, L., & Etchells, M. (2016). Assessing the academic, social, and language production outcomes of English language learners engaged in peer tutoring: A systematic review. *Education and Treatment of Children*, 39(3), 359-388.

Casallas, U.D.F. & Samir, F.P.C. (2016). Argumentation skills: a peer assessment approach to discussion in the EFL classroom. *PROFILE Issues in Teachers' Professional Development*, 18(2), 111-123.

Cheng, W., & Warren, M. (1999). Peer and teacher assessment of the oral and written tasks of a group project. *Assessment & Evaluation in Higher Education,* 24(3), 301-314.

Cheng, W., & Warren, M. (2005). Peer assessment of language proficiency. *Language Testing*, 22(1), 93-121.

Colby, D.C. (2011). Using *"Assessment for Learning" practices with pre-university level ESL students: A mixed methods study of teacher and student performance and beliefs.* ProQuest Dissertations Publishing.

Farrokhi, F., Esfandiari, R., & Schaefer, E. (2012). A many-facet rasch measurement of differential raters severity/leniency in three types of assessment. *JALT Journal*, 34(1), 79-102.

Güler, C. (2017). Use of WhatsApp in higher education: What's up with assessing peers anonymously? *Journal of Educational Computing research,* 55(2), 272-289.

Iraji, H.R., Enayat, M.J., & Momeni, M. (2016). The effects of self- and peer-assessment on Iranian EFL learners' argumentative writing performance. *Theory and Practice in Language Studies,* 6(4), 716-722.

Jafarpur, A. (1991). Can naïve EFL learners estimate their own proficiency? *Evaluation and Research in Education*, 5, 145-157.

Jingjing, X., & Fengying, M. (2015). A study into students' views on guided peer feedback in group work. *Chinese Journal of Applied Linguistics*, 38(2), 188-203.

Jones, I., & Alcock, L. (2014). Peer assessment without assessment criteria. *Studies in Higher Education*, 39(10), 1774-1787.

Linacre, J.M. (2002). What do infit and outfit, mean-square and standardized mean? *Rasch Measurement Transactions*, 16(2), 878. Retrieved from www.rasch.org/rmt/glossary.htm.

Linacre, J.M. (2012). Many-facet Rasch measurement: Facets tutorial. Retrieved from http://winsteps.com/tutorials.htm.

Linacre, J.M. (2018). FACETS: computer program for many faceted Rasch measurement (version 3.80.4). Chicago, IL: Mesa Press.

Matsuno, S. (2009). Self, peer, and teacher-assessments in Japanese university EFL writing classrooms. *Language Testing*, 26(1), 075-100.

Pandadero, E. & Alquassab, M. (2019) An empirical review of anonymity effects in peer assessment, peer review, peer evaluation and peer grading. *Assessment & Evaluation in Higher Education*, 44, 1253-1278.

Patri, M. (2002). The influence of peer feedback on self- and peer-assessment of oral skills. *Language Testing*, 19, 109-131.

Saito, H. (2008). EFL classroom peer assessment: Training effects on rating and commenting. *Language Testing*, 25(4), 553-581.

Saito, H., & Fujita, T. (2004). Characteristics and user acceptance of peer rating in EFL writing classrooms. *Language Teaching*, 8, 31-54.

Shin, R. (2013). Effect of using Facebook to assist English for business communication course instruction. *The Turkish Online Journal of Educational Technology*, 12(1), 52-59.

Topping, K.J. (2009). Peer assessment. *Theory in Practice*, 48(1), 20-17.

Vanderhoven, E., Raes, A., Montrieux, H., Rotsaert, T., & Schellens, T. (2015). What if pupils can assess their peers anonymously? A quasi-experimental study. *Computers & Education*, 81, 123-132.

Submission Guidelines

SUBMISSION GUIDELINES

Manuscripts for the JACET International Convention Selected Papers (JACET Selected Papers) will only be accepted by online submission. Please read the following guidelines carefully.

Important Dates:

Submission Form Open: October 7, 2019

Submission Deadline: 11:59 PM Japan Standard Time, January 14, 2020

Submission Form:

http://www.jacet.org/selected-papers-submission2019/

A. Requirements

1. A paper must be based on a presentation (oral or regular poster presentation) given at the JACET 57th International Convention and the first contributor must be a member of JACET. All other contributors must have also presented the work at the JACET 57th International Convention.
2. A paper based on a plenary lecture may be submitted as an Invited Paper (by invitation only).

B. Editorial Policy

1. JACET Selected Papers, a refereed, open-access electronic journal, encourages submission of the following:
   - -Research Articles on pedagogy and topics of significance to teachers of English
   - -Symposium Papers on relevant issues to teachers of English (one per symposium)
   - -Practitioner Reports to share findings and insights
2. Manuscripts submitted to JACET Selected Papers must not have been previously published, nor should they be under consideration for publication elsewhere.
3. Manuscripts which do not conform to the guidelines will not be considered for review.
4. Only one paper can be submitted by each contributor.
5. The Editorial Board of JACET Selected Papers reserves the right to make editorial changes in any manuscript accepted for publication to enhance clarity or style. The corresponding author will be consulted if the changes are substantial.
6. Paper offprints will not be provided.

C. Guidelines

1. Manuscripts on A4 paper, including abstract, references, figures, tables, and appendix, should not exceed 30 pages for Research Articles, 20 pages for Symposium Papers, and 15 pages for Practitioner Reports.
2. All manuscripts must be in English.

3. All submissions to JACET Selected Papers must conform to the requirements of the Publication Manual of the American Psychological Association, 6thedition.
   3.1 Prepare manuscripts using Microsoft Word (2003 or later).
   3.2 Use 12-point Times New Roman font.
   3.3 Leave margins of 2.5 cm on all sides of every page (A4 size, 210 mm × 297 mm or 8.27 in × 11.7 in). There are 26 lines to a page.
   3.4 Do not justify right margins.
   3.5 Do not use running heads.
   3.6 Create a paper without the author name(s).
   3.7 Include the title, an abstract (no more than 200 words), and keywords (no more than five keywords; a multiple-word key phrase can be counted as one keyword).
   3.8 Acknowledgements should not be included at the time of submission.
   3.9 For pagination, use Arabic numerals placed in the upper right-hand corner of each page.
   3.10 Throughout the entire manuscript including references, both author name(s) and their publication information should be substituted with "*****."

D. Submission Procedure

1. All contributors must complete a submission form on the JACET website, which can be accessed from the Submission Guidelines.
2. Contributors must follow the instructions below.
   2.1 Save the manuscript under the author's full name as in the following examples: suzukikaoru or smithkerry. Submit both Word and PDF files.
   2.2 If there are more than four authors, write all authors' information on a separate file and send to the JACET office by e-mail: Author names, affiliation, membership number, postal code, address, telephone number, and e-mail address.
   (JACET e-mail address: jacet@zb3.so-net.ne.jp)
   2.3 Do not include a cover sheet.

E. Contributor's Responsibility

1. Contributors are responsible for the content of their manuscripts.
2. Contributors are responsible for obtaining permission to reproduce any material such as figures and tables for which they do not own the copyright, and for ensuring that the appropriate acknowledgements are included in their manuscript.

F. Copyright

1. JACET holds the copyright of the articles published in JACET Selected Papers.
2. Anyone, including the author(s), who wishes to reproduce or republish an article, must obtain permission from JACET. Also, it should be clearly stated that JACET holds the copyright.


JACET