

A Letter from the Editor

Friends and Colleagues,

Greetings! We hope this issue finds you well. As the fall/winter term draws to a close, we hope you have plenty of reading time. This is our biggest and best issue yet, covering a wide range of topics.

As the quality and quantity of submissions continues to improve, we decided to print four feature articles this time. First, **Andrew Obermeier** describes the challenges and rewards of setting up a university-wide English curriculum. Next, **Christine Winskowski** presents *Part 2* of her study on student ratings instruments, offering practical advice and examples to help you custom-design SRIs to suit your needs. The third feature article is a study by **Lorraine Sorrell**, in which she compares the editing practices of international students before and after entering university. Finally, **Renée Sawazaki** presents her experience implementing a theme-based English curriculum at a Japanese university.

In *Research Digest*, **Christopher Long** provides an overview of the matched-guise technique for investing attitudes, and **Kaoru Kobayashi** contrasts the way metatext is used to maintain coherence in Japanese and English research articles.

In *Opinion & Perspective* **David Peaty** offers a spirited rebuttal to Michael Guest's critique which appeared in the last issue.

Rounding out this issue is one conference review and two teaching tips from the *Chalkface*. First, **David Ockert** and **Jerry Talandis, Jr.** report on JALT 2005, sharing the highlights of the conference held in Shizuoka last October. In *From the Chalkface*, **Nic Farrow** explains how he uses consciousness-raising to teach the articles, and **Steven Newman** shares his method of using conferencing as a warm-up activity.

As you are enjoying these articles, we hope you will consider becoming more actively involved with your publication. We have worked hard to improve the quality of this journal, and more improvements are coming soon. But we are a small staff, and we need help. *Professional Development*, *Cyberpipeline*, and *Book Reviews* all require editors. Proof readers are also needed. If you can spare a couple of hours a month, you can help us maintain this positive momentum. Contact me to volunteer.

And as always, we need your submissions. Whether you are an inexperienced writer or a widely published author, write something up, send it to us, and allow your name to grace the pages of *On CUE*.

We hope you enjoy 14.1.

Mike Hood
On CUE Editor

From the Editor

Constraints, Practices, and Needs in a University-Wide English Curriculum

Andrew Obermeier
Kyoto University of Education

京都教育大学のFD委員会による統計(平成16年度授業アンケート報告書)を分析することにより、本学の学生が全学英語科目に対して他の授業と比べて満足度が低いことが分かる。ここ20年に渡って、外国語必修科目が激減して来たが、これを見直す大学側の提案により2006年に科目数が久しぶりに増加する。本論文では、この絶好の機会に、本学の外国語英語科目の授業改善方針を論じ、組織性を高めるため、教科書の選択システムを提案する。

Introduction

In an era of fierce public scrutiny and competition for applicants, universities recognize that foreign language requirements must justify their contributions to overall institutional aims through proficiency gains that will be useful to students after graduation. Offering all students on one campus a series of English courses that serve their needs is not easily achieved when the curriculum serves students of widely varied majors. Learners' prior study experiences and interests vary vastly, so determining objectives, content, materials, teaching methods and assessment for courses can be difficult. Although teachers are often left to their own devices to decide how best to address these issues, institutions trying to make program-wide improvements need to find ways to encourage and assist teachers to collaborate toward achieving program-wide goals.

The University-wide English Curriculum (UEC) at Kyoto University of Education serves about 270 first-year students who select English classes to fulfill their foreign language requirement each year. All students must choose a foreign language to meet the university's general education requirement. This is a teacher's training university; students major in subject areas they intend to teach after they graduate, such as Math, Art, Science, Japanese, Sociology, and Physical Education. Although all students are required to take at least four foreign language courses, university leaders doubt the need for foreign language skills for schoolteachers. Major requirements and teacher licensing requirements limit student choices, restrict timetables, and regularly take priority over foreign language education. Thus the administration has steadily cut the foreign language requirement, shrinking the UEC from sixteen credits in 1983 to four credits today.

Recently, however, the Ministry of Education's new emphasis on English for international communication has resulted in a reversal of this trend. The local Board of Education's use of TOEIC scores as a qualifying credential for hiring elementary school teachers has revitalized the recognition of the importance of English study. The university administration added two courses to the foreign language requirement, starting in April 2006. This article will discuss the

processes involved in resuscitating the UEC amid this new enthusiasm for developing an effective curriculum.

The Current UEC

The UEC is now a brief series of English courses that students take to fulfill a four-course foreign language requirement. First-year students take *English Communication 1* and *English Literacy 1* in the spring, and *English Communication 2* and *English Literacy 2* in the fall. There are fifteen sections of each *Communication* class, with an average of 18 students per class. Ten sections of each *Literacy* class are offered, averaging twenty-seven students per class.

Three questions from class evaluations administered and statistically analyzed by the university's Faculty Development Committee reveal that the majority of students are adequately satisfied with their English

courses, but less so than they are with their other courses. In response to question 2, "How satisfied were you with this course overall?" 68.5% of students responded positively about UEC courses on the Likert scale questionnaire, compared to 78.9% that responded positively about their other university courses. The percentage of students that gave strongly dissatisfied responses concerning their UEC courses was disturbing, however. Table 1 shows responses to three questions from the March 2004 student evaluation.

Students submitted three hundred eighty-four course evaluations for UEC classes, filling out one for their *Literacy* class, and one for their *Communication* class. Three questions were selected from the evaluations to estimate students' interest (Question 1); their satisfaction (Question 2); and their impression of the teacher's passion for his/her subject (Question 3). Responses to Questions

Table 1

Comparison Between Responses to Selected Questions on Class Evaluations for UEC and All Other Classes

Response	UEC students (n = 384)	All others (n = 4638)
<i>Q. 1: Did you become interested in the course's theme?</i>		
Very interested	22.7%	23.7%
Moderately interested	44.0%	53.5%
Not so interested	21.1%	17.2%
Completely uninterested	12.2%	5.6%
<i>Q. 2: How satisfied were you with this class?</i>		
Completely satisfied	23.7%	26.5%
Mostly satisfied	44.8%	52.4%
Unsatisfied	17.4%	15.0%
Completely unsatisfied	14.1%	6.2%
<i>Q. 3: Regarding the teacher's passion for his or her subject:</i>		
Great passion	30.0%	39.5%
Some passion	48.6%	49.8%
Little passion	13.8%	8.7%
Absolutely no passion	7.6%	2.0%

1 and 2 showed that over twice as high a percentage of students responded that they were "completely uninterested in" or "completely dissatisfied by" UEC courses than they were for their other courses. In Question 3, a percentage more than three times higher felt that the UEC teachers had "absolutely no passion" for the course they were teaching. Question 2, perhaps the best measure of overall feelings toward courses, showed that 14.1% of students were completely dissatisfied with their UEC courses. Contrasted with only 6.1% who gave this

harsh judgment to their other university classes, this is an alarmingly high level of dissatisfaction. Since students responding so critically tend to be the most vocal, and since UEC classes are taken by 85% of all first year students, these results stand out for administrators. Student dissatisfaction, coupled with the local school board making English proficiency a necessary credential, has made curricular improvement a focal point for the university.

Curricular Processes

For the UEC, curricular developers will need to establish practices that bring long-term improvement. A theoretical framework suggesting processes for developing purposeful, accountable curricula is outlined in Johnson's (1989) multi-stage framework, which links the activities of decision makers at each stage of curricular development with specific contributions to the curriculum. Brown (1995) emphasizes assessment as essential for insuring that teachers and learners focus on achieving course objectives; tests are used to both understand learner achievement and evaluate program effectiveness. Nunan (1988) diverges from this systematization, warning that it should not be assumed that what gets planned will be taught

or that what gets taught will be learned. He emphasizes a research perspective toward the curriculum that continually evaluates what learners actually learn, and whether the outcomes of teaching are useful to them in their lives outside of the classroom.

At the outset of designing a new curriculum, some degree of top-down systemization and decision-making is inevitable, but processes should be established which enable teachers to better understand learner needs and teach responsively. In a coherent curriculum, the institution's purposes are immediately apparent in the classroom. Tajino, James, and Kijima (2005), considering a Japanese university setting, proposed systemic processes of course design for balancing learners' needs with institutional goals. They gave an example of a class in which the teacher defined shared purposes from what have long been oppositional ideological stances in English language education in Japan: English language learning for literacy cultivation and English language education for practical use.

Balancing between institutional leadership and teacher and learner autonomy is one of the most complex aspects of curricular development. Chain-of-command organizational structure may be cooperated with officially,

Table 2

Stages, Decision-making Roles, and Products in the UEC

Developmental stages	Decision-Making roles	Products
1. curriculum policy making	university administrators	policy documents
2. needs analysis means specification	UEC administrators	course objectives
3. course design materials selection	UEC administrators & teachers	syllabi teaching materials
4. Classroom implementation	teachers & learners	teaching acts learning outcomes

but will be resisted covertly in classrooms. Complete lack of organization is also unacceptable. Teachers need clear curricular goals and objectives, but they also need to be invested in achieving them. In Table 2, Johnson's (1989) model helps to outline how curricular processes connect to each other, and the different roles that faculty involved might fulfill. The model below is adapted so the roles assigned match the actual positions of people involved in the UEC.

The left column defines activities at each developmental stage, the middle column shows which faculty members are involved in the activity, and the right column specifies the outcomes. The outcome of the activities at each stage of development are called "products". The products of the different stages of development are the policy document, course objectives, syllabuses, teaching materials, lesson plans, teaching acts, and learning outcomes. Each product is incorporated into the activities at other stages of development in a recursive system. In a robust, and coherent curriculum, learning outcomes are carefully evaluated, and administrators consider them in future policymaking. Policy decisions are then integrated into the activities at the other stages of development. In curriculums where there is little coherence, classroom implementation happens haphazardly, without any program-wide consideration of the upper-level planning stages. The more teachers working in the lower stages on the chart are able to participate and influence the upper stages, the more coherent the curriculum becomes. That is to say, if teachers are closely attuned to learners' needs, and are able to give them voice in consultation with UEC administrators (or better, by fulfilling UEC administrative roles), then course objectives will be stated in such a way that materials selection and course design responds more

closely to learner needs.

This model, however, is probably a better fit for a commercial foreign language institute, where purposes can be specified for a calculated target market and students enroll on their own volition. The focus is diverted from the UEC, however, where teacher licensing requirements and major requirements compete for students' attention. Also, at commercial schools, faculty members are hired full-time to work within one curriculum, so this becomes the main content of their work. In the UEC, staff attention is divided away from it. Full-time faculty members are burdened with unrelated university and departmental duties. Part-time faculty members only come to the university twice weekly and spend little time at the university outside of class. For the UEC, sustainable policies will be those that accommodate faculty time constraints while focusing efforts toward matching course objectives to learner needs.

Due to the constraints against collaboration and extensive attention to the UEC, some degree of chain-of-command style leadership must prevail. There is simply no time for all teachers to gather and make large-scale decisions. For *Literacy* classes, for example, UEC administrators have decided to implement university-wide grammar midterms and finals. This should help to bring focus to course contents, but it might also instill an oppressive atmosphere. Hopefully, all teachers will be able to influence test design. In tandem with such a policy, careful diagnostic placement testing would also be helpful. If the placement test can give a good sense of proficiencies that most students have trouble with, and the midterm and final can urge learners towards learning such problem areas, then the testing policies might be of educational value. Another way to connect

learner needs to course objectives and syllabuses might be through close regular attention to student course evaluations and teacher reflection questionnaires. Yet another might be through guiding the collaborative selection of course materials, which will be outlined in the next section.

Materials Selection & Development

For *Communication* classes in the UEC, textbook choice is done individually, with teachers choosing texts to suit their teaching style and students. While such a situation is very practical for teachers, it results in a disparate curriculum, with parallel classes following separate texts and course outlines. Student complaints often address the widely disparate course designs and requirements of classes under the same title in the UEC. For *Literacy* classes, on the other hand, UEC administrators have chosen a textbook of grammar exercises and explanations. While this unifies course contents across the curriculum, the policy disgruntles teachers who resent being forced to work with a textbook they had no part in choosing. In the UEC, the dilemma between requiring a text across the curriculum and allowing teachers free choice might be avoided if administrators provide teachers with clear course objectives and encourage them to collaborate in textbook selection. Teachers should be given flexibility concerning the day-to-day contents of their courses, but this should be within the bounds of a consensus decision about textbook adoption.

Chambers (1997) outlines a process for choosing textbooks by prioritizing desirable textbook features as they collaborate to decide on a common text. Collaborating to decide on texts might be a meaningful exercise in deepening understanding of curricular aims and clarifying objectives. A

group consensus about a textbook is a shared statement of commitment, as well as a contribution to the contents of the curriculum. While mandatory collaboration might seem stifling, a happy medium should be found for materials selection that allows for both flexibility and unity of purpose. As Allwright (1981) asserts, textbooks offer teachers prepared options for achieving a given set of objectives and assist them in their roles as managers of learning. Masuhara (1998) suggests that since the elements of course design are now built into ready-made textbooks, selecting a textbook frees teachers to focus on delivering instruction by giving a framework for course structure. If a group of teachers decides to use the same textbook, they establish a common ground for discussion about the course. Teachers need not feel required to cover every textbook activity, they should view the textbook as a set of ready-made teaching options. Indeed, teachers can choose to use a text book only as a source of teacher reference, using it to loosely guide topic choice, target functions, and activities. Teachers should be encouraged to supplement lessons with their own materials as needed, and a file of supplementary materials for each class could be made that teachers could use as a resource to expand or personalize their lessons.

The Overall Purpose of the UEC

Before textbooks can be selected, however, the overall purpose of the UEC must be clearly stated. The central pillar of curricular development is responsiveness to learner needs, and curricular theorists consistently prescribe needs analysis as the essential step in determining curricular goals (Brindley, 1989; Brown, 1995; Richards, 2001). UEC students' learning needs, however, cannot be

easily categorized. Students in the UEC have a wide variety of majors, and as a whole will use English in uncountable ways. Further, the field of needs analysis itself is complex: learner's desires, skills lacked, proficiency standards, necessary pragmatic competencies, and suitable learning styles are just a few of the perspectives from which a needs analysis can be made. Therefore, the UEC strives toward the all-encompassing aim of "English for General Purposes," and proficiency standards based on the Test of English for International Communication (TOEIC) will be implemented. Since TOEIC scores serve as qualifications in many different fields in society, and increasingly for public school teachers as well, aiming instruction at score improvement appears to be a simple answer to the broad range of student needs, and a practical way to prepare students for the workplace.

Basing university instruction on practice for proficiency tests is problematic, though, being a fundamental abdication of educational responsibility. The TOEIC and other proficiency tests are designed to measure overall English proficiency, and not to guide instruction. These tests cover a broad range of target grammar structure, vocabulary, and usage, and present them within a wide variety of contents. Teaching to improve test scores will therefore be guesswork unless it focuses on practicing test-taking strategies. Although the English department is inclined against it pedagogically, the clear, concrete goal is very alluring from an administrative standpoint. Despite the fact that no one can grasp the contents that need to be taught to raise TOEIC scores, and test preparation drill work does not seem like an appropriate university course topic, focusing on the TOEIC gives students, teachers, and future employers a clear, measurable definition of

ability to aim toward. Reluctantly, UEC administrators have decided to implement this benchmark.

Curricular development is often an exercise in compromise, and this is also true in establishing central purposes. "General use," and "intercultural understanding" are too broad to be meaningful goals. Students have already studied toward similar purposes in junior and senior high school. The MEXT guidelines place an emphasis on face to face, spoken communication. At the university level, students are more focused than they were in secondary education, and pursue academic development in a chosen field. While the stated aims of secondary foreign language education are appropriate for the beginning stages of foreign language learning, they are too vague and commonplace for university learners. Deeper, more involved explorations using English academic skills would enhance student abilities further. Students are now focusing on academic skills in their other classes, so their English classes should congruently focus on these skills.

Cummins' (1976) review of research on second language acquisition and cognitive growth argued that bilingual development is a process in which learners first develop "Basic Interpersonal Communication Skills" (BICS), and then progress to "Cognitive Academic Language Proficiency" (CALP). BICS are characterized by social communication skills that are context embedded and cognitively undemanding. CALP requires the ability to understand academic language that is not supported by the rich array of non-verbal and contextual clues that characterize face-to-face interaction.

While Japanese secondary students are appropriately developing highly contextual communication skills, university educators would do better to aim curriculums at help-

ing students to express and understand more abstract concepts. Perhaps students could, through English, present, write about, and discuss their academic interests in a seminar-like setting in English classes. A shift in emphasis of English in higher education toward more academic uses might help the overall paradigm by raising expectation levels and expanding course contents.

Conclusion

This paper has explored issues involved in curricular development at a small Japanese university. Administrators and teachers need to focus on learner needs and respond to them when they make policies and course objectives. It has shown how Johnson's (1989) framework for curricular development plays out in a Japanese university, where teachers and administrators have diverging priorities. The limited opportunities for teacher collaboration and sustained attention to developing courses make UEC administration difficult. In all, developing a coherent curriculum at even a very small Japanese university entails institutional leadership that can compromise between institutional ideals, professional practices, and classroom realities.

References

- Allwright, R. (1981). What do we want teaching materials for? *English Language Teaching Journal*, 36(1), 5-18.
- Brindley, G. (1989). The role of needs analysis in adult ESL programme design. In R. K. Johnson (Ed.), *The Second Language Curriculum* (pp. 63-81). Cambridge, UK: Cambridge University Press.
- Brown, J. D. (1995). *The elements of language curriculum*. Boston, MA: Heinle and Heinle.
- Cummins, J. (1976). The influence of bilingualism on cognitive growth: A synthesis of research findings and explanatory hypotheses. *Working Papers on Bilingualism*, 19, 1-43.
- Chambers, F. (1997). Seeking consensus in coursebook evaluation. *English Language Teaching Journal*, 51(1), 29-35.
- Johnson, R. K. (1989). *The second language curriculum*. New York: Cambridge University Press.
- Masuhara, H. (1998). What do teachers really want from course books? In B. Tomlinson (Ed.), *Materials development in language teaching* (pp. 239-260). Cambridge: Cambridge University Press.
- Nunan, D. (1988). *The learner centered curriculum*. Cambridge, UK: Cambridge University Press.
- Richards, J. (2001). *Curricular development in language teaching*. Cambridge, UK: Cambridge University Press.
- 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*, 3, 13-22.

Documenting Instructor-Effectiveness, Part 2: Toward Low-Inference Student Ratings Instruments

*Christine Winskowski
Morioka Junior College
Iwate Prefectural University*

本論は、一般的に行なわれている高推論型の学生による授業評価法 (SRI) の短所を論じたパート1の続編である。パート2では、SRIの質問項目に対していかに多様な解釈が可能であり、学生の偏った考察が評価結果に影響を及ぼすかを示してSRIの実効性を検証する。この問題はしかし、教師自身の手で授業目的に合わせた低推論型のSRIを展開することによって解決が可能である。比較的具體性の高い設問は推論への依存度が低く、授業内容に対する学生の直接的観察と学習内容に即した回答が期待できる。本論では様々な低推論型のSRI設問モデルを提示する。個々の教師の必要に応じてこれらのモデルをアレンジすることが可能であろう。

After the semester is over, I approach my course evaluations with a mixture of eagerness and trepidation. I suppose that I am not the only teacher who flips quickly through the conventional student ratings instruments (SRIs) to see if responses are *generally* on the positive side of the scales. Then I check the comments—anything overtly critical? I doubt I am entirely alone in wincing mentally if some of my ratings are not on the high end. As I slow down and take a more detailed look, I dare say I am like most people, wondering what made a particular student choose this response or make that comment.

I feel less trepidation when I am checking out responses to SRIs that I have designed for my own classes, in contrast to those required

in my university. The reason is that I gain a lot more genuine information about how students were engaged with and affected by the course from my own SRIs than from the university's. Whether a student gives a positive or negative rating on an item, the rating is less vague and affective, and more specific and meaningful (at least to me).

The difference in substance between the two types of SRIs has motivated this paper and its precursor, *Part 1: Vulnerabilities of Conventional Student Ratings Instruments*. In *Part 1*, I argued that conventional SRIs suffer a variety of weaknesses, including the frequent absence of formal validation procedures (no one knows how frequent), inadvertent influences from the organization, the form of the items themselves, demonstrable response-biases, and questionable discrimination of teaching effectiveness from ineffectiveness. In this paper, I argue that instructor-designed SRIs with low-inference items, based on course objectives and tailored to the events of the course, provide much more authentic, substantive, and useful information than conventional SRIs for understanding of a particular instructors' effectiveness.

High- & Low-Inference Items

In the field of education, the concepts of high-inference measurement and low-inference measurement are well-known, although they are regarded as relative, rather than ab-

solute, terms.¹ *High-inference* measures (in this case, rating systems) are conceptual and abstract, attempting to capture a quality in a person or event. For example, ratings items that ask if the instructor is “willing to help students,” “respects students as persons,” or is a “fair” grader, are relatively high-inference items (these examples are taken from Benz & Blatt, 1996). In other words, these items ask student raters to *infer* an instructor quality that cannot be directly observed. High-inference items seem to correspond to what may be regarded as a conceptual definition (Bernard, 2000) of the “teacher effectiveness” variable. However, items which ask about particular behaviors that can be more directly observed, such as whether the instructor starts class on time, speaks “clearly and audibly” (Benz & Blatt, 1996), or regularly invites questions and comments from students, are relatively *low-inference* items. Low-inference items may contribute to an operational definition (Bernard, 2000) of “teacher effectiveness.” Items asking students to observe their own behavior, or the events of the course, could also be considered low-inference (assuming they are well-constructed; more on this below).

Student Rater Cognition and High- & Low-Inference Items

As argued in *Part I*, conventional SRIs tend to have many high-inference items, which define teaching effectiveness in a conceptual (rather than operational) way (Winskowski, 2005a). Such SRIs may also presuppose a particular teaching paradigm. Kolitch and Dean (1999) examined the 22-item omnibus student evaluation instrument used at their institution, noting that it is “typical” of those used in the U.S. The authors wanted to determine what sort of teaching paradigm is reflected in its items.

They concluded that their instrument presupposes a teacher-centered transmission model of teaching in which knowledge is discrete and measurable, and students acquire this knowledge from the teacher and master it, according to predetermined objectives.² If this is typical of student ratings forms (e.g., those exemplified in *Part I*), then regardless of what kind of course is actually being rated, students are responding to a set of implied roles—teacher as transmitter and themselves as relatively passive receivers of knowledge. Of course, if this model of teacher-centered transmission does not fit the course design, the value of students’ ratings must certainly be questioned.

Furthermore, not only do students make inferences about the qualities of the instructor or course events in conventional SRIs, but when the instructor reviews the SRIs after the end of the course, the instructor in turn must infer what students might have been thinking, reasoning, and referring to as they filled them out!

Recent research suggests that if we look beyond student ratings numbers, there is a lot to learn about what actually goes on with students as they fill out their SRIs. In fact, Burden (2005) pointed out in a recent issue of *OnCUE* that we know little about the process inherent in students’ responding to evaluation forms. While his study focused on students’ understanding of the uses to which SRI are put and their opinions on the matter, a smattering of studies has begun to make clear that students’ interpretation of SRI items and reasoning behind SRI response selection is not uniform, even if their responses appear to be. At the same time, these studies raise the issue of whether students’ cognitive processes differ when responding to high-inference and low-inference items.

To wit, as noted in *Part I*, a recent study

by Billings-Gagliardi, Barrett, and Mazor (2004) used thinkaloud interviews of students as they completed a conventional SRI which addressed the design of the course, the course materials and methods, and the instructors' facilitation of student learning. The authors found that students defined terminology key to their field differently from one another (sometimes idiosyncratically); they interpreted the meaning of response options (strongly agree, agree, disagree, etc.) differently from each other, and made selections that may have been biased by how they anticipated the evaluations would be used; and they made judgments about teacher effectiveness based on criteria that sometimes had nothing to do with teaching effectiveness. The findings of this study came out of the examination of students' responses to the SRI as a whole, rather than any part or item.

In another study, Kolitch and Dean (1998) focused not on the entire SRI at their institution, but on the one global item taken as the "best indicator of an instructor's effectiveness" (p. 120), item No. 22 "Overall, [the instructor] was an effective teacher," a relatively high-inference item. Kolitch and Dean asked 96 students to write down their "conceptions of a 'good' teacher" and their views on SRIs" (p. 122). They also followed up with semi-structured interviews with a representative subset of students. Like Billings-Gagliardi et al., they found students had multiple interpretations of item 22: Some equated "overall" with "on the average" and made their assessment by averaging positive and negative aspects of the instructor's performance. Other students focused on a dimension of effectiveness that they felt was critical to effectiveness (e.g., "enthusiasm," "approachability," "challenge to students"), and based their assessment on that. Still

others fell back on personal feelings about the instructor, as was found with some in the Billings-Gagliardi et al. (2004) study. Moreover, students reported confounding influences on their evaluations, including criticism about items and cynicism about the value of SRIs, as well as fear of reprisal. (The latter two findings were also observed among some students in Burden's 2005 study and reflect biases unrelated to teaching effectiveness).

Similar findings were also found by Benz and Blatt (1996), who examined responses for each of eight instructor-rating items. They showed that students gave multiple interpretations to SRI items, offered a variety of evidence for their ratings, had different assumptions about the teaching process, and indicated ambiguous and contradictory thinking about their own ratings.

Also in this study, the authors noted that across the responses for these items, the variability in student response was *not* related to whether items were high-inference or low-inference. For example, an item in Benz and Blatt's study about the instructor being "well-prepared" was seen by various students as subject expertise, punctuality, and readiness with overheads. An item on the instructor's "clarity of presentation" was met with the most variable responses (including understanding, sequencing of material, and personalizing of the material). These both are relatively high-inference items, and a variety of interpretations could reasonably be expected. However, a relatively low-inference item on the "clarity and audibility of the instructor's speech" had two interpretations: diction and vocabulary. At the same time, some relatively high-inference items on qualities of being "respectful," "fair," and "helpful" toward students generated stronger consensus (shown in students'

synonyms for these qualities) than some low-inference behavioral items in Benz and Blatt's research. The authors noted that an item about whether the instructor "presented material in an interesting way" evoked the strongest consensus in students' interpretation, namely "interesting" meant the use of story-telling. The authors concluded that high or low levels of item-inference seemed unrelated to the degree of agreement students showed in their rendering of item meaning (Benz & Blatt, 1996).

This may seem a counter-intuitive finding. Does it mean that low-inference items give us responses that are no more concrete than high-inference items? Possibly in some cases, and possibly not in others. A complete picture of what kinds of items prompt multiple response types or instill a sense of uncertainty or ambiguity, etc. in student raters must wait for a consensus of research. And if a modest prediction may be permitted here, that consensus will emerge. As Schwarz (1999) has noted, responding to an SRI tacitly follows the conventions of discourse, and as we know from a wealth of conversation analysis and discourse analysis, discourse follows scripts and is highly patterned. It is entirely reasonable that either a high- or low-inference item could evoke a single meaning or multiple meanings which are consistent across populations with particular statistical probability.

What Do Low-Inference SRI Items Tell Us?

So if students might have multiple interpretations of low-inference SRI items, why would such items be better and more informative than high-inference items? The remainder of this article illustrates the value of low-inference item content and design.

To accomplish this, I would like to broaden

the meaning of the term *low-inference item*. Much of the literature addresses low- and high-inference items in terms of students' perceptions of how the instructors do their job. Here I propose expanding the focus to include items on how students do their job, how the design of the course does its job, and how the synergy of the students' engagement, course design, and instructor effectiveness combine and play out in the context of the course objectives. Actually, this is the focus of those "nosey" questions we ask ourselves about what happened in class, how things worked, how students found this element and that activity, etc. To illustrate, here are some nosey questions I found myself musing on last semester (with answers from my own, low-inference SRIs):

- *In the English listening/speaking class, how did students find the change from a standard 90-minute period to two 45-minute periods per week?* (Seems that most liked it.)
- *Do the students feel they are processing and dealing with more English language than before?* (It's not clear.)
- *Do they feel they are learning more English?* (Yes, I'm relieved to say.)
- *Do they like the new textbook better than the old one?* (Comments suggest they did.)
- *In the lecture class comparing Japanese and U.S. cultures, did answering the study questions help students sort out what parts of the lecture were most important?* (Most thought so.)
- *Did providing a lecture outline help them follow the flow of the lecture?* (Seems like it did.)
- *Was I sensitive enough in noticing when the English I used in lecture was unclear or too complicated for my Japanese*

students? (Japanese students' reticence to ask questions or complain is well-known, so who knows? But almost half said they had questions left unasked on the weekly study sheet provided, so probably I wasn't.)

- *Do some students have complaints about my courses?* (One person said the speaking class is too easy. Some said the lecture class was too hard; a few said they would not recommend it to others, though most would.)
- *Did I do anything annoying or aggravating or frustrating to the students?* (Someone told me I should write on the board more "prettily.")
- *Did they appreciate the value of the on-line component and how it complemented the on-site component?* (I didn't think to ask this.)
- *What about the glossaries in the new on-line component—were they helpful?* (It seems like they didn't think so, but I noticed that the web site logs showed students didn't access them much.)
- *Why didn't students use the glossaries more?* (No idea.)

These illustrate questions an instructor might have about actual events of the course with their actual participants. In contrast to the rather bloodless and occasionally mystifying items one might find in conventional SRIs ("The instructor conveyed the points clearly." "The instructor used the text and materials well."), low-inference items that would answer the "nosey" questions above offer substantive and less equivocal data about what really happened in a course. Furthermore, they offer concrete information about what improvements are needed and how they might be implemented.

Readers will note that answers to several

of my nose questions were gleaned from low-inference SRI that I have designed for my classes. (I know, for example, that I need to re-think the mechanism for students to ask questions and for use of the on-line glossaries.) The next section addresses the nature of low-inference item construction.

What Makes Low-Inference SRI Items Low-Inference?

Low-inference SRI items differ from high-inference items in two ways: design and content. Design of low-inference items moves away from the rather minimalist structure that we often see in conventional SRIs. As illustrated in *Part 1*, raters are most commonly presented with choices on a "5- or 6-point Likert-type scale indicating a continuum, from agreement to disagreement, high (excellent) to low (poor) value, high to low frequency, or something similar" (Winkowski, 2005a). These are then paired with a series of statements which are to be rated on the scale.

Making the scales themselves more detailed and descriptive is one way to lower the inference level of an item. For example, of all the commercial SRIs discussed in *Part 1*, Kansas State University's IDEA form has the best approximation of low-inference rating scales and items in its section on students' rating of their own progress: "Using a scale of 1 – lowest 10% of courses I have taken, 2 – next 20%, 3 – middle 40%, 4 – next 20%, and 5 – highest 10%, several items ask students to indicate their progress on gaining factual knowledge, learning fundamental principles, developing specific skills and competencies, developing oral and written skills, etc." (Centra, 1993, p. 185). Note, however, that the principles, skills, etc. are not specified.

Felder (1993) also advocates specific rat-

ings scales for course and instructor ratings which reduce the amount of inference needed to understand the rating. He offers this illustration and comment:

“Rate the instruction you received in this course on a scale from 1 to 5, with 5 being the highest response.” Ratings of this sort are most effective when the numbers on the response scale are clearly defined. [Rather than] “excellent,” “above average,” “fair,” etc., ...you can get greater discrimination with a variation of the following instruction: *When responding, use as a basis of comparison all of your previous high school and college teachers. A response of 5 denotes one of the three or four best you’ve ever had; 4 = top 25%; 3 = 40-75%; 2 = bottom 40%; and 1 = one of the three or four worst you have ever had.* An instructor whose average rating is close to 5 on this scale is clearly doing a superb job and deserves nomination for an outstanding teacher award, and *serious* problems obviously exist if an instructor’s rating is consistently close to 1. Ratings close to 4 indicate commendable teaching performance and ratings close to 2 suggest the need for corrective measures. (Felder, 1993, p. 1)

While this remains a global teacher effectiveness item as found in most conventional SRIs, the specific definition of the choices would go a long way toward forestalling inconsistent rater-interpretation of choices and the need to infer what students meant when they made their ratings. Other design alternatives for low-inference items are illustrated in the following section.

Low-inference items are also distinguished

by their specific content. Armstrong (1998) asks, “Why not assess learning directly? Ask the students to answer questions about their own performance because it is the students, not teachers, who are the producers of learning” (p. 1223). He suggests that students be asked if they were clear about objectives; if they were prepared and organized; if they spent much time learning, did their best, and what it was that they learned. (Actually Armstrong appears to be assessing learning processes here, rather than learning itself.) Hake (2002a, 2002b) echoes this notion in his advocacy of assessing the cognitive impact of a course—the success of the learning—rather than the affective impact.

Felder adds, “Ask open-ended questions on mid course evaluations, leaving plenty of space for the responses: 1. *What do you like best about this course and/or the instructor? (List up to three things.)* 2. *What do you like least about the course and/or instructor? (List up to three things.)* 3. *If you were the instructor, what would you do to improve the course?*” (Felder, 1993, p. 2).

One of the few other authors advocating instructor-written assessment items is W.L. Rando (2001). He describes four kinds of instructor-written items:

- Open-ended, self-report questions for exploring students’ experiences (e.g., “How did analyzing case studies help you prepare for the final project?”)
- Closed-ended, self-report questions for identifying specific aspects of students’ learning experiences (e.g., on a 5-point scale from Definitely Yes to Definitely No, “Analyzing cases taught me to apply theory to practice.”)
- Open-ended, direct assessment questions for exploring what students are learning in a general way (e.g., “Apply Theory

A to the following situation, and write a paragraph on your conclusions.”)

- Closed-ended direct assessment questions to identify specific areas of understanding, confusion and learning (e.g., “According to theory A, which of the following is true about the paragraph you just read? (a) Workers are unmotivated. (b) Workers are under prepared. (c) Workers are over qualified.”). (Rando, 2001, pp. 78-79; examples from p.80)

Rando goes on to point out that standard SRI items are frequently too vague (i.e., high-inference) to produce useful results. He offers three pieces of cogent advice. First, instructor-designed items should be based on specific objectives for the course. “Learning objectives are precise statements of the changes we are trying to create in our students. When we ask ourselves how we expect students to be different as a result of our teaching, we can construct questions that truly speak to our core intentions and then assess the effects of the methods we choose” (p. 79-80). Next, Rando advises asking “questions that are reflective opportunities for students” (p. 81). Tailoring an item about how an assignment changed students’ thinking trains students to witness their own learning processes and become better learners as a result. Finally, Rando encourages faculty to ask specific questions of students about which elements of the course teaching methods further the students’ learning and fit the course and departmental objectives, and which do not (Rando, 2001). The next section offers some alternatives based on Rando’s advice.

What Do Low-Inference SRI Items Look Like?

Below are several illustrations (adapted

and expanded from an earlier version in Winskowski, 2005b) which may serve as templates or examples for low-inference item-types. These are all taken or adapted from my courses, and many are content-specific. Instructors can tailor these items, adapting and combining the templates to configure a ratings instrument that reflects their own intentions and objectives for student learning, course events and activities, and students’ engagement.

Ratings Of Skill Change Through The Course With Specific Scale Descriptors

This template requires that instructors do two things. First, identify specific descriptors of skill, ability, or knowledge level, following the advice of Felder (1993) and Bain (1996). Instead of the usual vague “very poor - poor - average - good - excellent” descriptors, make specific descriptors to quantify the skill or ability level. For example:

- 1 [My skill increase was] none at all
- 2 some, but not enough for this class (or some, but only enough for previous courses in this course sequence)
- 3 enough for this class
- 4 enough for other classes in my department (college, university, etc.)
- 5 enough for professional ability in jobs

Second, as Rando suggests, isolate five or six (or more) key outcomes—skills, abilities, areas of knowledge—that students should leave a course with. One or more statements for each outcome must then be generated that allow students to indicate how their skills and abilities had changed from the beginning to the end of the course, using a form like the following:

My ability to (or awareness of, skill at,

understanding of, etc.) XXX went from _____ to _____.

Of course, ratings scales and outcomes should be tailored to the discipline and the course topic. For example, the scale may detail levels of recognition and understanding of historical processes, skill in mathematics or language, problem-solving ability in engineering, applications in social science or education, or critical thinking in any field.

Here are some examples from courses I have taught in composition and English as a second language. The key outcomes should be apparent from the statements; the rating scales approximated the model indicated above.

- a) Students in a writing class could show, using the appropriate number:

My ability to organize my thoughts in writing went from (e.g., 2) to (e.g., 3).

My ability to organize a paragraph went from ____ to ____.

My control of grammar in writing went from ____ to ____.

- b) In a listening class, students could tell me:

My ability to understand the main idea of a lecture went from ____ to ____.

My ability to take effective lecture notes went from ____ to ____.

My recognition of academic vocabulary went from ____ to ____.

- c) In a reading class, students could indicate:

My understanding of paragraph organization went from ____ to ____.

My ability to predict what comes next in reading went from ____ to ____.

My reading speed went from ____ to ____.

This type of evaluation item can also be

used to monitor program-wide skill and ability development. For example, in my department, we have used a larger array of items naming 21 English language skills to monitor the effects of our language curriculum. It is administered when each new class enters the university and again at graduation. (See Winskowski & Hanna, 2004, for a preliminary report.)

Descriptive Clusters For Each Course

Activity/skill Area

This type of item also has two elements—identification of an activity, goal, or event, and the ways in which the students might have responded. Statements can be written describing the range of responses to the activity and placed in clusters. Then students can simply check off what is true of their own case. A fill-in alternative may also be provided for atypical student experiences. Here are some examples from an in-depth unit on library skills in a composition course:

- a)

____ *I knew most of the information in the exercises.*

____ *I knew some of the information in the exercises, but learned a lot about library systems from doing the exercises.*

____ *a good deal of the information in the library workbook was new to me.*

- b)

____ *Now I feel confident about finding information from reference books, indexes, and periodicals*

____ *Now I'm still uncertain whether I can successfully find information from reference books, indexes, and periodicals*

____ *Other – please explain* _____

In the following case, I wanted to check my perception of students' participation in an opportunity to ask questions in writing after every lecture:

- c)
- ____ *I sometimes had questions about the lecture, and wrote them on the Study Question Sheet.*
- ____ *I sometimes had questions about the lecture, but did not write them on the Study Question Sheet.*
- ____ *I never had questions about the lecture.*

From this cluster, I discovered what I suspected was true: Almost half of the class responded to the second choice, showing they had had questions, but had not asked them. When students experience some element of a course that changes, a similar set of alternatives could be offered. The following were asked after some changes were made in an English language class after the first year of a two-year sequence.

- d)
- I'm speaking English in class _____ more than last year.*
- I'm speaking English in class _____ the same amount as last year.*
- I'm speaking English in class _____ less than last year.*
- e)
- The English I'm learning is _____ more than last year.*
- The English I'm learning is _____ the same as last year.*

The English I'm learning is _____ less than last year.

f)

The new English textbook is _____ more useful than last year's book.

The new English textbook is _____ the same as last year's book.

The new English textbook is _____ less useful than last year's book.

Finally, value can always be added to an item by following it with "because _____."

Ranking Course Activities

If there are several course activities that share a general learning objective, students can be asked to rank them for their effectiveness and explain why. For example, if reading-response paragraphs, classroom discussion, and case studies are expected to contribute to critical thinking skills, a rank-order item may be presented this way. This case is an example of *forced ranking* (i.e., ranked elements are forced into an order from high to low):

a)

Please rank order the following items to show what was most effective in promoting your critical thinking skills. *1 = most effective; 3 = least effective. Explain, in each case, why.*

____ *Written responses to reading, because _____*

____ *Classroom discussion, because _____*

____ *Case study analysis, because _____*

A disadvantage of forced-ranking is that we cannot be sure whether students think the bottom-ranked activity is really *ineffective* or simply *less effective* than the top-ranked items. An alternative is to list all of the course

activities or group them according to skill, objective, etc., and ask students to give a *criterion-based ranking* (e.g., how helpful they were in contributing to learning, how useful they were in achieving a goal, etc). Below is an example taken from a foreign language listening and speaking class, where essentially all activities are expected to contribute to the constellation of sub-skills which make up listening and speaking ability. The example below refers to several activities of the class.

b)

Please put a number by the following class activities to show how helpful for learning English each one is.

1 = very helpful, 2 = somewhat helpful, 3 = not so helpful

___ *watching the video text episodes on my own*

___ *watching video text episodes in class with some explanation from the teacher*

___ *exercises from the video book*

___ *exercises from the conversation book for two people*

___ *exercises from the conversation book for three-four people*

___ *practicing dialogue from the video episodes (first reading, then saying)*

___ *learning about collocations from the story dialogue*

___ *exercises on pronunciation, intonation, and word stress*

___ *exercises from supplementary conversation text (the green book)*

___ *audiotaped conversation homework*

___ *email conversations with classmates*

___ *other (please explain)*

Please explain what was especially helpful about your top 5 choices.

The following examples focus on an on-line component of an on-site course.

c)

Please put the number that is true for you about the course website.

1 = very useful, 2 = somewhat useful, 3 = not very useful, 4 = I didn't use it

___ *The course website (in general)*

___ *The weekly glossary*

___ *The lecture outlines*

___ *The resources (web sites, newspaper articles, etc.)*

___ *The forum (threaded) discussions*

Note that students' perspective on what are helpful activities may be quite different from what an instructor imagines, as my colleagues and I have sometimes found out.

Short-Answer Items

An informal approach to inviting student feedback is to simply pose questions about aspects of the course, and have students respond with short answers, following Felder's (1993) and Rando's (2001) advice. With this type of item, instructors can get rapid feedback at any point in the semester, can investigate whether there are trouble spots and what their nature is, and deal with them in a timely fashion. Here is an example taken at mid-semester from a linguistics course:

a)

- *Are you learning a lot of linguistics?*
- *What is the most interesting topic in the course? The least interesting topic?*
- *What do you feel will be most valuable in this course to you for your major?*

- *How do you like the research project you are working on?*
- *What has been the most difficult part of the course so far?*
- *How is the lecture format (i.e., Power-Point notes) working for you?*
- *Would you suggest any changes for the remainder of the semester?*
- *Etc.*

This sort of item has endless possibilities. It could also be useful if the instructor has implemented a new component in the course, or changed an element mid-term. Is a new on-line component functioning in the way that the instructor intends? Is student access to a resource (like a lab) working in the way it should? Have the students succeeded with some out-of-class phase of work? Or short-answer items can simply be used for general feedback.

Routine Feedback

In a course setting that regularly involves a pattern of activity (e.g., new lecture material or new lab project each week or in each unit), students can be trained to provide regular feedback. Here is a simple example of a brief feedback that would work for lectures:

- a)
- What percent of the lecture did you understand? ____%*
- Is there a part of the lecture that should be briefly re-explained? Please describe: _____*
- What questions do you have? _____*

The instructor must, of course, determine whether and how to respond to this feedback. Should a special session be held for a small group of students who seemed to

have missed a critical item? Should a point be reviewed for the whole class? And so on. Additionally, feedback items can be written to ask about specific lecture content (or the content of another routine activity, such as laboratory work).

Soliciting this kind of feedback on a regular basis does require the instructor to be flexible with the classroom schedule in order to accommodate unexpected problems or issues that come up.

Conventional-Style Ratings With Specifically Tailored Items

Finally, as Bain (1996) suggests, the conventional format of student ratings forms can be combined with specifically-tailored outcomes of a course. A standard five-point scale (strongly agree – agree – not sure – disagree – strongly disagree) can be used with items that are written for the outcomes of a particular course. Following are examples for two classes, one comparing Japan and the U.S. and the other on business English. This type of item is quite useful for directly asking if, in the students' experience, course objectives were met.

- a)
- ____ *I understand U.S. American culture better after being in this course.*
- ____ *This class made me see the Japan differently from before.*
- ____ *When I completed the study questions, they helped me to understand what was important.*
- ____ *Now I understand better the similarities and differences of Japan and the U.S.*
- b)
- ____ *I learned how to construct English-style business letters.*
- ____ *I learned how to construct English-*

style business memos.

— *The textbook was useful for colloquial business English expressions.*

— *I learned some practical knowledge about writing (e.g., about punctuation, capitalizing, etc.).*

— *I regularly used the listening tapes that went with the textbook.*

The inference-level of any one of these items could be made still lower. For example, instead of “I learned how to construct English-style memos,” an instructor could ask students whether they feel they could construct a memo with “Full confidence and ease,” “Some confidence, relying on models,” or “No confidence” as is done with items in section 2 above.

It can be seen that these items are intended to focus on the events and activities of the course, how the student was engaged with them, and what the student took away from them—largely what cannot be known in much detail from conventional SRIs. While it must be conceded that some of these items involve degrees of judgment, many involve students’ direct reporting of their experience in the course, and are relatively more objective, requiring less inference, than might be found in conventional SRIs. Direct inquiries about the value of activities of the class work well for this format.

Because the instructor can tailor items that refer authentically to specific events of the course, the face validity of such items is far higher than many found in conventional SRIs. This in turn lays the foundation for a ratings instrument with greater content validity and conceptual validity. Ultimately, such an instrument is simply more informative and substantive for evaluating the instructor’s effectiveness at designing and delivering the course.

Conclusion

The admonition to complement SRIs with other forms of faculty assessment is near universal, as those familiar with the 2000 or so pieces of this very large body of research literature will attest. However, conventional SRIs remain the single most-used form of faculty assessment. Peer-observation, faculty portfolios, interviewing of student representatives, etc. are more cumbersome to arrange, implement and analyze and are less “objective” appearing—and, one must suspect, comparatively rare.

The low-inference style of SRI that is suggested here provides a wealth of information that conventional instruments cannot. Further, students can observe that this low-inference style of evaluation item comprises genuine questions about the effectiveness of various elements of the course. Students who become accustomed to this type of evaluation will likely become more discriminating learners, more conscious of both their own effort in a class and the value of various parts of the class to their learning. The key is the authenticity and relevance of the items with regard to the instructor’s intentions and the students’ experience.

Since the use of student ratings instruments is in its infancy in Japan, there is an opportunity to pioneer a new approach to the very real need for sensible and meaningful evaluation of courses and instructors. Conventional SRIs can continue to play their role where they must, generating data that allows some comparison across equivalent kinds of classes, and across time, for example. However, low-inference, instructor-designed alternative SRIs can provide the more immediate, substantive, and meaningful data that complements—or replaces—conventional SRIs.

References

- Armstrong, J.S. (1998). Are student ratings of instruction useful? *American Psychologist*, 53, 1223-1224.
- Bain, K.R. (1996). *Using student ratings*. Retrieved March 15, 2005, from New York University Center for Teaching Excellence Web site: <http://www.nyu.edu/cte/white.html#Number2>
- Benz, C.R., & Blatt, S.J. (1996). Meanings underlying student ratings of faculty. *Review of Higher Education*, 19(4), 411-433.
- Bernard, H.R. (2000). *Social Research Methods*. Thousand Oaks: Sage.
- Billings-Gagliardi, S., Barrett, S.V., & Mazor, K.M. (2004). Interpreting course evaluation results: Insight from think aloud interviews with medical students. *Medical Education*, 38(10), 1061-1070.
- Burden, P. (2005). Student ratings: Meaningful voices or meaningless ritual? *On CUE*, 13(2), 16-25.
- Centra, J.A. (1993). *Reflective faculty evaluation*. San Francisco: Jossey-Bass.
- Felder, R. (Winter 1993). What do they know anyway? 2. Making evaluations effective. *Chemical Engineering Education*, 27(1), 28-29.
- Hake, R.R. (2002a). Lessons from the physics education reform effort. *Conservation Ecology*, 5(2), 28. Retrieved March 15, 2005, from <http://www.ecologyandsociety.org/vol5/iss2/art28/>
- Hake, R.R. (2002b). Problems with student evaluations: Is assessment the remedy? Message posted to POD listserv, archived at <http://www.stu.ca/~hunt/hake.htm>
- Kolitch, E., & Dean, A.V. (1998). Item 22, "Overall, [the Instructor] was an effective teacher": Multiple meanings and confounding influences. *Journal on Excellence in College Teaching*, 9(2), 119-140.
- Kolitch, E., & Dean, A.V. (1999). Student ratings of instructors in the USA: Hidden assumptions and missing conceptions about 'good' teaching. *Studies in Higher education*, 24(1), 27-42.
- Rando, W.L. (2001). Writing teaching assessment questions for precision and reflection. In K.G. Lewis (Ed.). *Techniques and Strategies for Interpreting Student Evaluations*. New Directions for Teaching and Learning. San Francisco: Jossey-Bass.
- Schwarz, N. (1999). Self-reports: How the questions shape the answers. *American Psychologist*, 54(2), 93-105.
- Winskowski, C. (2005a). Documenting instructor-effectiveness, Part 1: Vulnerabilities of conventional student ratings instruments. *On CUE*, 13(2), 2-15.
- Winskowski, C. (2005b). Student ratings of faculty and courses: Investigating instructor effectiveness. *Bulletin of Morioka Junior College*, 7, 39-62.
- Winskowski, C., & Hanna, C. (2004). Junior college students' perceptions of English skill improvement. *Language and Culture*, 6, 61-78.

Footnotes

¹ Definitions of these terms are offered in Rosenshine & Furst's book *Research on teacher performance criteria* (1971), cited in R.C.Chavez (1984), "Use of high-inference measures to study classroom climates: A review," in *Review of educational research*, 54(2) 237-261.

² The authors contrast his paradigm to one in which the course also incorporates student knowledge and experience, and in which students are engaged as co-creators of the curriculum, its objectives, etc.

Beyond Correction: A Comparison of Editing Practices of International Students Before and After Entering University

Lorraine Sorrell
Macquarie University

本論は、オーストラリアの大学で大学進学前の留学生と学部生である留学生の文章における誤りとその自己修正の数について比較を行ったものである。大学進学前の学生に対するIELTS(アイエルツ)テスト対策クラス及び学部生の授業の中から小論文の草稿を取り上げ、文法・綴り等の誤りについて調査し、誤りの内容及び第一稿と比較した第二稿の誤り修正数について分析を行った。小論文の第二稿では、学部生のグループはIELTS対策クラスのグループと比較して格段に低い誤りの再発率を示した。学部生の別グループに対して行われた小論文のフォローアップ調査では、若干の自己発見法の要素が取り入れられたが、誤り修正の比率は元のグループと殆ど同一であった。

Introduction

The inspiration for this study arose when I noticed that my pre-university students seemed to take more care with editing their writing than my undergraduate students did. International students studying in English-speaking universities are frequently told before entering university that fluency in speaking and writing is more important than accuracy, and that it is more important to convey the meaning than to be too concerned about correct grammar, vocabulary or mechanics. This is, of course, true, especially when they are low-level users of the language, and when the first drafts of a paper or an essay are being planned. It does sometimes seem, however, that they have not learnt that the situation

is different at university, where academic writing requires mastery of the *conventions* of English writing. Perhaps they have learnt the lesson of *fluency before accuracy* almost too well. As an English language teacher and a university lecturer, I can see that while the teaching and encouraging of fluency is sound in every way, students going on to overseas English-speaking universities need more than fluency. One of the many challenges students need to address is adjusting to, and becoming familiar with, *all* the requirements of the new discourse communities (Johns, 1998; Swales, 1990) or *academic discourses* (Gee, 1996) they find themselves in.

Gee (1996) argues that what is important is “language *plus* being the ‘right’ *who* doing the ‘right’ *what*. What is important is not language, and surely not grammar, but *saying (writing)-doing-being-valuing-believing combinations*” (p. 127). Although Gee explicitly excludes grammar, it is arguable that attention to form, of which grammar is a part, in academic writing is *included* in that combination of saying, writing, doing, being, valuing and believing. As Swan (1999) notes, echoing Eskey (1983) and Coe (1987), acceptable writing indicates that writers have gained access to and belong to the discourse community of choice. They then know how to conform to the norms of that particular society. This is what Swan is referring to when he says “we have to live in the world as it is”

(p.62). Violation of academic expectations, or nonconformance, can “mean a forfeit”, writes Purves (1991, p.36). The forfeit, in the case of international university students, can mean the misinterpretation of their academic ability, lower marks, and in serious cases, the loss of opportunities to engage in higher research degrees.

Ferris (2002) notes that “accuracy is important in the real world to which student writers go [and] a lack of accuracy may both interfere with the comprehensibility of their message (or ideas) and mark them as inadequate users of the language” (p.9). Song and Caruso (1996), refer to claims that “ESL students are at a disadvantage in situations where their essays are evaluated by instructors without ESL training or where they are expected to have language proficiency equal to that of native students” (p.164). Dickins and Woods (1988) discuss a number of research findings that show that learners with poor grammar are more likely to underachieve academically.

Kaldor, Herriman, and Rochecouste (1998) note that in Australian universities it seems that international students are particularly likely to have their academic competency underestimated because of their “language behaviour and cultural attitudes” (p.viii). Bush (1994), in her survey of academics’ attitudes to writing, found that 76% strongly agreed that grammatical accuracy was important and “80% strongly agreed with the statement *Poor grammar distracts the reader*” (p. 21). In a follow-up to this survey, Desierto (1998) found similar results, where 64%-84% of academics thought that accuracy of grammar is important in writing, and 64%-89% agreed that poor grammar distracts them from understanding the content. Inaccurate vocabulary was also listed as a source of concern for 85.4% of the respondents.

Comments such as the above are confirmed by the findings from a survey at Macquarie University (Jones & Roger, 2000) which indicate that in many departments there is some disquiet about the formal writing skills of students. Statements such as the following verify this disquiet:

- *A student with sloppy grammar is demonstrating a lack of concern for detail which is almost invariably related to lack of concern for, or knowledge of, other principles of logical argument... sloppiness in grammar goes with sloppiness everywhere.*
- *Main problems seem to be sentence structure and grammar. As a result, arguments and content don't come across clearly.*
- *Lack of skills (grammar and spelling) commonly makes communication impossible.*
- *I marked “grammar” first because an incoherent and grammatically incorrect paper can hardly present a coherent or convincing argument.*
- *If failure is the outcome of poor language skills, then so be it.*

Although there is evidence in the literature that there are academics who are prepared to give qualified acceptance to problems with form from non-native English speakers (see for example Bush, 1994; Janopoulos, 1992; Vann, Meyer & Lorenz, 1984), overall, research, anecdotal evidence and my own observations make it clear that there are many academic instructors who are seriously worried about the accuracy of the writing of their students.

This lack of accuracy can interfere with meaning (Eskey, 1983). Ferris (2002) quotes numerous writers (for example, Janopoulos, 1992; Santos, 1988; Vann, Meyer & Lorenz, 1984) who have noted that in university set-

tings errors are “distracting and stigmatizing” (p. 9). However, there is little research literature about the differences in error correction behavior between pre-university and university students.

This study was motivated by curiosity about different attitudes I had noticed in my classroom. While engaged in individual writing consultations with international university students from all faculties, I frequently noticed negative comments from lecturers about surface level grammatical errors in assignments. In most cases these grammatical errors did not affect the meaning of the texts, but they were indicative of an incomplete mastery of academic use of the English language.

Most comments explicitly stated that it was the language problems that were preventing the students from gaining higher marks in their assignments. It was clear that many lecturers at university expected a higher level of linguistic accuracy. The study was thus motivated by two factors: the constant evidence of comments from lecturers from different disciplines in the university; and my own observations of the differences between my two types of classes, which were both pre-university and undergraduate.

The study examined one aspect only of second language writing; the end of the writing process—editing. The word *editing* here is used in Hyland’s (2002) sense of the “final stage in the writing process where the writer attends to surface-level corrections of grammar and spelling” (p.230). Polio, Fleck, and Leder (1998) also define editing as relating to sentence-level changes during revision, with the conscious application of explicit knowledge.

Setting

The small, classroom-based study was based at an Australian university, and the English Language Centre attached to it. At least 20% of the students at this university are international students. The classes investigated were IELTS preparation classes at the English Language College (pre-university), and EAP (English for Specific Purposes) undergraduate classes at university. All participants were from Asian countries; a large percentage were Japanese.

An important detail is that although EAP stands for *English for Academic Purposes*, at this university it is not simply an English Language course. The students had already passed the IELTS examination, and the EAP course here taught academic reading and writing skills (with the emphasis on writing) and academic cultural awareness.

The instructional focus in the classes attended by both sets of participants included a strong emphasis on content and structure in writing, but less emphasis on form, or surface errors. The overarching pedagogical position was that teaching grammar was not the answer to the “problem.” All the students in this study had studied English formally for at least six years, and had either achieved a score of 6 or more on the IELTS (International English Language Testing System) examination or equivalent, or were close to doing so. Thus the assumption was that they “knew” grammar, but in fact, they often could not, or did not, apply this knowledge in their academic writing.

The aim of the study was to compare the correction rate of surface writing errors. The participants were given no explicit correction of those errors, but they were given dedicated time to revise. The project looked at genuine student-produced documents, written in authentic classroom settings. We felt that this

was as informative as examining research reports (Eskey, 1983), and potentially more useful than examining fabricated errors and corrections in decontextualized documents.

The research question was: In academic writing, what are the observable differences in correction rates of surface errors between two different types of classes: a pre-sessional IELTS preparation class, and an EAP undergraduate class.

First and second drafts of essays were examined for surface errors. A total of 82 essays in the first study and 22 in the follow-up study were analyzed for surface errors. In line with the current philosophy of the English Language teaching sections of the university, there was no explicit teaching of grammar in class before writing, nor was there any explicit feedback on the grammar in the first drafts of the essays. The aim was to find out what the process of rewriting and self-correction would achieve.

The participants were chosen from two different classes. Group One was composed of 15 pre-university students from an IELTS Preparation class at the English Language College attached to the university. These students were preparing for the IELTS test in order to gain access to university, and had entered the preparation class with the equivalent of 5.5 on the IELTS Band Score.² They were aiming for a Band Score of 6.0 and were studying English exclusively and intensively, with 20 contact hours a week.

Group Two was composed of 26 undergraduate students who had already started their first semester at university. Their English proficiency was variable, but as they had needed a Band Score of 6.0 to enter the university, they all had a minimum competence of Band 6.0. These students came from a variety of disciplines but were all studying EAP (English for Academic Purposes).

EAP is a credit-bearing undergraduate unit specifically restricted to students with a non-English speaking background.

A significant difference between the EAP students and the IELTS preparation students was that the undergraduates were studying EAP as one of four academic subjects in a thirteen-week semester, and had only four contact hours per week (lectures and tutorials), compared to twenty contact hours per week in the IELTS class.

Process

Throughout the course of the study, as little disruption as possible was made to the normal running of the classroom. It was regarded as crucial for all data to be obtained in as natural a situation as possible, for the sake of future use of the findings. All research activities were normal classroom activities and were not modified in any way. First and second drafts of essays from both classes were examined. The IELTS class essays were quite short, averaging 300 words. The EAP essays were considerably longer, from about 900 to 1200 words. See Appendix A for the essay questions.

In the IELTS Preparation class, the essays were written as part of the normal classroom program, with the first draft being written as homework. The second draft was written in class. The EAP first drafts were normal university assignments, written at home, and were worth 30% of the final mark at the end of the semester. The students had previously submitted a plan for the essay, when structure and content had been discussed and feedback given. Students in both classes wrote their first draft and submitted it for marking. As was the practice, they were marked for content and structure, but surface errors were not marked or located. Before the students rewrote the essays, the teacher discussed in

both classes the most frequent surface errors, and gave mini-lessons on how to revise them. The students were then asked to rewrite the essays, focusing on all aspects of content, structure *and* error correction.

The IELTS students re-wrote their essays in class. The EAP students had two weeks at home to rewrite their essays, but of course at that time they were also writing other assignments and studying for final exams. All students were encouraged to use dictionaries and grammar books when revising their essays. The first and second essay drafts were then photocopied, marked, and categorized for surface errors. The errors were then noted on a separate error form for each student.

The method of error analysis chosen was adapted from Kroll (1990, p. 143), who wrote a descriptive analysis of 100 essays written by undergraduate non-native speaker university students. Kroll made a sentence-level syntactic analysis of her corpus of essays, and identified 33 different categories of error. The error form used here was adapted from Kroll (1990, p.145) and Ferris and Roberts (2001, p.169), and contained fourteen categories of errors (See Appendix B).

To compensate for the considerable difference in the length of the two different class essays, an error/word ratio was used. A standard measure to form a basis for an accuracy ratio is to use the total number of words in a text and tabulate the number of

errors (Kroll 1990, p. 146). The error/word ratio was counted in both essay drafts, and the percentages of errors in both were calculated. The percentage of improvement was then calculated. Table 1 gives an example of how the error corrections were calculated.

The differences in the error correction rate between the two groups was quite large, with the IELTS class achieving an average error correction rate of 34.3% compared to the EAP class which achieved an average rate of 20.2% - a difference of 14.1%.

This higher IELTS correction rate was interesting as, according to Kroll (1990) writing under pressure of time (i.e., in class) is not likely to result in the students' best work. Would the correction rates have been even higher if they had done their revisions at home instead of in the classroom? The lower correction rates of the EAP students suggest that this might not have been the case. The results can be directly compared with Kroll (1990) and Polio et al. (1998). As in these studies, both the IELTS and the EAP classes made improvements in their second drafts. In Kroll's study, the students who rewrote at home had no higher correction rates than those who rewrote in class. This is comparable to the results here, although the EAP students, who rewrote at home, actually had a *lower* correction rate than the IELTS students, who rewrote their essays in class.

There are several possible explanations for the much higher rate of corrections in the IELTS preparation class. These students were still English Language students and had an important English exam to pass in order to enter university. Also the essays were revised in class, under supervision. Perhaps this element of com-

Table 1
Error Calculations

First draft	Words	Errors	Percentage
Student X	311	50	16.08%
Second draft			
Student X	297	22	7.41%
			Change
			-46.08%

pulsion and the dedicated time for re-writing motivated them to work harder to correct their errors than if they had rewritten the essays at home. These students had less time to rewrite their essays than the EAP students but, arguably, more motivation.

The students in the EAP class were much busier, as they were also studying three other subjects. As they had already passed their English examinations and entered university, it is possible that they felt that they no longer needed to concentrate on their English language. In short, the EAP students had more time to rewrite their essays than the IELTS students, but perhaps had less motivation.

The strong contrast in error correction rates between the two class types led to the planning of a further study with EAP students the following semester. It was hoped that a more explicit awareness-raising component to classroom procedure would produce a higher correction rate in an EAP class, more comparable to the IELTS class. The same research procedure was followed, with small modifications to the teaching program, with a different EAP class.

EAP Follow-Up Study

The aim of the new study was to see if a small amount of consciousness-raising added to the normal classroom dialogue would lead to a measurable improvement in error correction rates. The same problems as with the first study existed: the necessity to fit in with a pre-arranged program; the fact that it was not an English Language course but a study skills course; and the fact the students were studying this course as just one fourth of their semester's work. The lectures could not be changed, but the tutorials did offer a little time for extra discussion about presentation and the need for final editing before handing in assignments.

Eleven students, mostly Japanese and Korean, in a new EAP class participated in this follow-up study. The new essay question is in Appendix A.

The following small changes were included in the tutorials:

- Short class discussions about the importance of final editing before submission were timetabled into the tutorials.
- Students were told about the previous study revealing the significant differences between the error correction rates of IELTS and EAP students.
- The students were told about the previous survey by Jones and Roger (2000), eliciting academics' opinions of inadequately edited work.
- Students were shown an overhead transparency of one comment from the Jones and Roger (2000) survey - "*A student with sloppy grammar is demonstrating a lack of concern for detail which is almost invariably related to lack of concern for, or knowledge of, other principles of logical argument...sloppiness in grammar goes with sloppiness everywhere*". The choice of that particular overhead was intended to "shock" the students, and it did indeed cause much amusement and concern and was the impetus for more discussion. One not entirely unexpected piece of news was that most students wrote their assignments the night and morning before they were due, and did not have the time, energy, or inclination to do any editing after that!
- A small amount of class time (one hour, half of one tutorial) was allocated to allow peer and self-editing in class, before taking the essay home to rewrite for the following week.
- Short consultations between students and the tutor about language use in the draft

essays were scheduled.

The results of this new study were somewhat disappointing. The objective had been to see if consciousness-raising would result in any improvement. However, in spite of the above changes, the error improvement was actually marginally worse than the original EAP study. This class had an average error correction rate of 19.3% compared to the previous EAP class's error correction rate of 20.2%. Table 3 shows the results of the three error correction studies.

Table 2
*Average Improvement Rate
of the 3 Groups*

IELTS	EAP 1	EAP2
34.3%	20.2%	19.3%

Polio, Fleck, and Leder's (1998) study showed a similar result with their experimental group, who did not show any significant improvement despite receiving extra help with editing. Although it was disappointing, there were some possible explanations for the lack of improvement in the new study:

- The essay question was generally regarded as being more difficult than the previous one, so it is possible that the students expended most of their energy on the content of the essay and had little time or energy left for editing.
- The participants were different students, with possibly completely different abilities.
- Although the tutor and the students discussed the *importance* of editing, and they were given dedicated class time to actually *do* editing, there was no time to actually teach them *how* to edit.

The results of the essay study show that with an error correction rate of only 20.2% (or 19.3%) in the EAP undergraduate classes,

compared to 34.3% in the IELTS pre-sessional class, there is clearly a problem with the EAP students' editing skills. Although the lack of an obvious improvement rate in the follow-up appears discouraging, it is not a reason to abandon the attempt to raise awareness and teach editing skills. The comments received from the students indicated that they felt the discussions had been valuable. Kroll (1990) suggests that it is possible that students simply do not know how to edit. However, although it seems reasonable and logical to teach editing skills, it is not actually known how useful this is likely to be, as there is little research examining the long-term effects of teaching editing skills (Ferris, 2002; Polio, Fleck, & Leder, 1998). But this should not preclude us from devoting more class time to explicit teaching of editing. Although the research on the efficacy of actual grammar instruction in classes is unclear and incomplete (Ferris, 2002), the external situation of disapproval from university lecturers still exists, so perhaps a stronger and more salient focus on editing, including the explicit teaching of editing skills, would be the next step.

It would be beneficial to:

- raise students' awareness of the negative impressions that 'bad' presentation creates in lecturers, with the resultant possible implications for grades;
- help them realize the possibility of written surface-level errors interfering with comprehensibility;
- give them practical instruction on how to edit.

Ferris (2002) points out that there are three reasons why students need to take editing seriously:

- Not all subject lecturers will ignore written language problems to focus on subject knowledge.

- In some universities there may be a “threshold of accuracy” (p.78) that students have to pass.
- Employers expect a certain standard of writing from university graduates.

It should be borne in mind that in this context the students’ difficulty was with the *application* of their grammatical knowledge, not with its acquisition. They need to be explicitly shown *all* the academic writing requirements and expectations involved in writing at university. University students need to develop both their fluency *and* their accuracy, and as Eskey (1983) notes, “Even in this age of ... coexisting with error, giving students what they need is still what good teaching is all about” (p.322). It is certainly not helpful for students to have lecturers so distracted by their linguistic errors that they cannot appreciate the content of their assignments.

It is hoped that teachers of Japanese students planning to enter an English speaking university will find this study useful or thought-provoking. Communicative language teaching is a valid and essential methodology but potential international university students need more specific information and instruction that will enable them to succeed in their new academic environments.

References

Bush, D. (1994). *Academic writing: faculty expectations and overseas student performance*. Paper presented at the National Conference of ACTA – WATESOL, Perth.

Coe, R. M. (1987). An apology for form: Or: who took the form out of process? *College English*, 49, 13-28.

Desierto, A. (1998). UWA expectations of academic writing at Australian universities: Work in progress. In B.Black, B. and N. Stanley, (Eds), *Teaching and learning in*

changing times (pp. 91-95). Proceedings of the 7th Annual Teaching Learning Forum, The University of Western Australia, February 1998. Perth: UWA. Retrieved February 2, 2004, from <http://cea.curtin.edu.au/tlf/tlf1998/desierto.html>

Dickins, P., & Woods, E. (1988). Some criteria for the development of communicative grammar tasks. *TESOL Quarterly* 22(4), 459-480.

Eskey, D. E. (1983). Meanwhile, back in the real world... Accuracy and fluency in second language teaching. *TESOL Quarterly*, 17, 315-323.

Ferris, D. (1995). Teaching ESL composition students to become independent self-editors. *TESOL Journal*, 4(4), 18-22.

Ferris, D. (2002). *Treatment of error in second language student writing*. Ann Arbor: University of Michigan Press.

Ferris, D., & Roberts, B. (2001). Error feedback in L2 writing classes: How explicit does it need to be? *Journal of Second Language Writing*, 10(3), 161-184.

Gee, J. P. (1996). *Social linguistics and literacies: Ideology in Discourses*. (2nd ed.). London: Falmer Press.

Hyland, K. (2002). *Teaching and researching writing*. Harlow: Pearson Education.

IDP Education Australia. (2003). *IELTS handbook*. Canberra: IDP Education Australia/IELTS Australia.

Janopoulos, M. (1992). University faculty tolerance of NS and NNS writing errors: A comparison. *Journal of Second Language Writing*, 1, 109-121.

Johns, A. M. (1998). *Text, role, and context: Developing academic literacies*. Cambridge: Cambridge University Press.

Jones, A., & Roger, P. (2000). [Attitudes of academics to students’ writing]. Unpublished raw data.

Kaldor, S., Herriman, M., & Rochecouste, J. (1998). *Tertiary student writing*. Sydney: NCELTR, in association with the University of Western Australia.

- Kroll, B. (1990). What does time buy? ESL student performances on home versus class compositions. In B. Kroll (Ed.), *Second language writing: Research insights for the classroom*. Cambridge: Cambridge University Press.
- Polio, C., Fleck, C., & Leder, N. (1998). "If I only had more time:" ESL learners' changes in linguistic accuracy on essay revisions. *Journal of Second Language Writing*, 7(1), 43-68.
- Purves, A. C. (1991). Clothing the emperor: Towards a framework relating to function and form in literacy. *Journal in Basic Writing*, 10, 33-53.
- Santos, T. A. (1988). Professors' reactions to the academic writing of non-native speaking students. *TESOL Quarterly*, 22, 69-90.
- Song, B., & Caruso, I. (1996). Do English and ESL faculty differ in evaluating the essays of native English-speaking and ESL students? *Journal of Second Language Writing*, 5(2), 163-182.
- Swales, J. (1990). *Genre analysis: English in academic and research settings*. Cambridge: Cambridge University Press.
- Swan, M. (1999). How much does correctness matter? In R. Berry, B. Asker, K. Hyland & M. Lam (Eds.), *Language analysis, description and pedagogy* (pp. 53-63). Hong Kong: Hong Kong University of Science and Technology.
- University of Sheffield English Teaching Centre (1998). *English language tests*. Retrieved October 31, 2003, from <http://www.shef.ac.uk/eltc/toefl98.html>
- Vann, R., Meyer, D., & Lorenz, F. (1984). Error gravity: a study of faculty opinion of ESL errors. *TESOL Quarterly*, 18, 427-480.

Appendix A

Essay question for the IELTS class: *Discuss the problems of overpopulation in the world, and provide some possible solutions that could be applied.*

Essay question for the EAP undergraduate class: *Academic writing is not uniform and cannot be learnt simply by studying "rules" for good writing. Students need to understand that writing occurs in a particular context, and that they may need to vary their writing to suit different purposes and different audiences. Discuss the factors which may lead to variation in academic writing.*

New essay question for the follow-up study in an EAP class: *"Globalisation ... is a complex set of processes, not a single one." Analyse some of the major processes or developments of recent decades, which can be included under the term "globalisation".*

Appendix B

Error Analysis Form

Class _____

Date _____

Student Name _____

Home Country _____

		Draft One	Draft Two
		Word Count	Word Count
Verb Error	verb form		
	verb tense		
	subject-verb agreement		
Noun Error	plural ending incorrect, omitted or unnecessary		
	possessive ending incorrect, omitted or unnecessary		
Article Error	article or other determiner incorrect, omitted or unnecessary		
Word Error	incorrect lexical choice		
	missing or unnecessary words		
	preposition error		
	pronoun error		
	spelling		
	word form		
Sentence Error	sentence/clause boundaries		
Punctuation	missing/unnecessary mark		
TOTAL			

Adapted from Kroll (1990, p. 145) and Ferris and Roberts (2001, p. 169).

(Footnotes)

- ¹ This was an informal survey of lecturers’ opinions, conducted as a needs analysis for the EAP section, at Macquarie University in 2000. The findings were not published.
- ² A band score of 5.5 falls between that of a ‘Modest User’ at Band 5 and a ‘Competent User’ at Band 6. At this university the normal IELTS requirement for undergraduate entry at that time was Band 6.00. (It is now Band 6.5). The TOEFL equivalents for IELTS 6.00 are 550 for the paper-based version and 213 for the computerized version (University of Sheffield English Teaching Centre, 1998).

A Thematic Approach in EFL Curriculum Design

Renée Sawazaki
Surugadai University

各課ごとに一定のテーマで編集されている総合教材を使い、日本人研究者がリーディングの部分を担当し、著者がコミュニケーションの部分を担当した。学生はある課のテーマに関する文章をリーディングの時間に読み、つぎに同じテーマについて著者とコミュニケーションの学習をする。著者らはこれをThematic Curriculumと呼び、実験群(TC)と対象群(non-TC)を比較することによって、その効果を測定した。その結果、語彙・文法、発表能力、学生アンケートなどいずれの項目においても実験群が有意の高得点を挙げた。この結果をもとに、複数の教師が同じクラスを各課ごとに一定のテーマで編集されている同じ総合教材を使って教えるカリキュラムの有効性を論じた。

Background

As a part of a university-funded research project at Surugadai University, two instructors, myself and Prof. Mieko Tsukamoto, in the Faculty of Cultural Information Resources created a thematic curriculum. We used the same course book and themes, such as learning English, vegetarianism, and care of the elderly, while focusing on the skills of reading in one course and oral communication in the other. The impetus for this study came from curriculum design in L1 and L2 education.

In the curriculum design practices in the disciplines of L1 learning, immersion bilingual education and foreign/second language education, we can find practices that include the use of content instruction as a vehicle to teach language. Integrated curriculum has been popular in American elementary and

junior high schools since the movement toward whole language instruction in the 1980s (Goodman, 1989; Grisham, 1995). This approach requires the primary instructor to use the same theme, for instance bears or the rain forest, to teach core subjects such as science, math and language arts. Research suggests that use of themes in written and oral language practice is beneficial for students' language skills development (Walmsley & Walp, 1990; Galda, 1998).

In the field of L2, such practices are often referred to as content-based instruction (CBI). The goals of CBI, as described by Wesche (1993) are for students to learn content and related second or foreign language skills concurrently, ultimately developing autonomy in using the target language.

Yet, in both of these practices, we see that it is the same teacher using a theme to teach knowledge and a variety of skills. In learning about practices in both of these disciplines, the instructors who conducted this study wondered if such practices could be transferred to a foreign language program where the same group of students take different skills-based foreign language courses taught by different instructors.

Strong (1995) documented a case of such a program. In the Integrated English program at Aoyama Gakuin University, teachers primarily used commercially published materials to teach combined skills, listening

and writing classes. He argued that using a theme-based curriculum promotes a movement from traditional grammar-based exercises into more communicative activities. In addition, his study suggested that themes can be highly motivating for both teachers and students since teachers can generate a wealth of materials and students can focus more on meaning rather than form.

In this paper, I will explore a smaller scale one-year program for non-English majors and focus on data collected at the end of the first semester. Our version of a thematic curriculum consisted of coordination of the syllabus for our respective classes to include the synchronized use of the same thematic units in a textbook. This structure was based on the following ideas concerning syllabus and curriculum design.

Many communicative foreign language textbooks are based on what may be called a topical syllabus. These textbooks include topics apparently selected on the basis of the authors' sense of relevance to the lives of the students for whom the text is designed (Brown, 1995). For this study, we considered textbooks that had a communicative focus yet included reading passages, such as *Impact Issues* (Day & Yamanaka, 1998) and *Impact Topics* (Day & Yamanaka, 1999).

First year students in our program are required to take two English courses: oral communication skills and reading skills. The teachers, most of whom are part-time instructors, are free to choose materials that match the goals of the course, yet they are often unaware of what is being taught in their students' other English class. Collaboration and use of topical themes across courses could create a link and add structure to the overall curriculum design.

One issue that we kept in mind in planning this study is that a thematic curriculum may

make teachers' workload easier and teaching experience better. The responsibility of individual teachers to create their own syllabus every year is quite burdensome. Literature in the 1980s stressed teacher involvement in curriculum renewal (Kouraogo, 1987) and more recently action research led by teachers has been encouraged (Hansen, 1997). Yet with the multitude of non-teaching duties of the average university teacher, it is not practical that all partake in every aspect of curriculum design. Focusing on a thematic curriculum may be a solution that would benefit teachers and students while satisfying the administration at the same time.

Brown (1995) divides curriculum design into four areas: approaches (ways of defining what and how the students need to learn), syllabuses (ways of organizing the course and materials), techniques (ways of presenting the materials and teaching) and exercises (ways of practicing what has been presented). As teaching is a deeply personal act, I suggest that when working for uniformity in programs, curriculum designers should focus on a thematic-based syllabus.

There is a web of related issues that can be explored. For the purposes of this small study, I focused on issues primarily related to how students reacted to using the same themes and course materials in their two required English courses.

Research Questions

1. What effect does a thematic curriculum have on student performance?
2. What are students' attitudes toward using the same themes in both Oral English and English Reading classes?
3. What are the benefits and challenges of creating a thematic curriculum for individual teachers?

Participants & Course Design

A one-year case study was conducted involving two first-year university English classes. The two collaborating teachers taught different skills to the thematic curriculum (TC) group. One section focused on reading skills while the other focused on oral communication skills. Both the control group and the TC group used the same materials for the oral communication section, and the TC group used the same content for the reading section. The control group, however, used different materials in the reading section, and the content of their two English classes was not integrated at all.

I taught the oral communication classes and will focus on data from these classes in this paper. Tsukamoto, the teacher for the reading section, published the results from these classes in *Surugadai University Studies* (Kiyomi et al., 2004).

These two courses were required for freshmen in the Faculty of Cultural Information Studies and met for 90 minutes once a week for two semesters. Students were grouped according to their freshmen seminar course, so the range of English proficiency was narrow. The average TOEFL score of the 267-student freshman class was 371, though not all students took the test. The TC group had 28 students and the control group 24. TC group students who took the TOEFL averaged 364, control group students, 384.

The TC group used the textbook, *Impact Issues: Thirty Key Issues to Help You Express Yourself in English* (Day & Yamanaka, 1998). The control group used this textbook in the oral English section, but not in the reading section. The textbook was chosen because it was suitable for a reading class as well as an oral communication class. Each unit consists of two pages. The first page contains a monologue or conversation illustrating one

person's experience with and opinion on a specific issue. The other page includes two sections: *What do you think?* in which five people give various opinions, and *Looking at the issue* which varies in style but focuses on expanding on the issue and elicits students' opinions and ideas. A section in the back of the book provides interaction tips with questions and phrases for expressing opinions. The student book includes a CD, with one story recorded at natural speed for each unit.

We selected 19 of the 30 units for the syllabus. As we needed to create the syllabus before the course started, we were unable to make selections based on students' preferences. In subsequent years, feedback from previous students could be taken into consideration. We did not know which issues would be particularly interesting to students, so we eliminated units based on our intuition. In the end, 12 units were used by both instructors.

For homework in the oral English section, students listened to and read the story, summarized it, wrote their opinion on the key issue and supported it with reasons. As the classes met only once a week, I regularly engaged students in a review of the story as a warm-up exercise in the following lesson. Various other tasks such as writing out mind-maps were assigned. The goal of the assignments was to help students prepare to express their ideas orally in class. In-class oral activities included group- and pair-work, teacher to student true/false or short answer questions, Think-Pair-Share (Kagen, 1992) and other cooperative learning tasks, and presentations. Consistent with the tenets of CBI, students were able to focus on sharing their ideas in English rather than on the language itself.

Method

Data

Three sets of data were collected to evaluate the effect of a thematic curriculum on student performance (Research question 1): (a) student written homework in the form of summaries of the stories; (b) a test on all the units that were covered in the spring semester; and (c) overall final grades. In addition, teacher observations on student ability to complete tasks were also considered.

To evaluate student attitudes and preferences (Research question 2), a questionnaire was given.

Finally, considerations for teachers implementing a thematic curriculum were based on reflections of the two teachers involved in this study.

Procedure

At the beginning of the fall term, a review test covering the six units that were studied in the spring semester was given. The units included: (a) Why Learn English?; (b) Forever Single; (c) What's for Dinner?; (d) Take the Money and Run; (e) For the Sake of the Children; and (f) Family Values. The test consisted of four parts: (a) vocabulary; (b) structures; (c) content; and (d) opinions. The test was written entirely in English and required both accuracy and fluency skills. I created the test myself, as a commercial one that included both language skills and content knowledge did not exist. Validity of the test is questionable, as it was not tested prior to this study.

The grading policy for each class was the same. A proportionate weight was given to written assignments, discussions, a group and individual presentation, test and final interview. This policy put more weight on active

engagement with the content of the course rather than on proficiency in English.

In order to assess student attitudes toward using thematic course materials in their English Reading and Oral English classes (Research question 2), a questionnaire was given at the end of the spring semester (see Appendix). Although the primary purpose was to gather information regarding the use of the same themes in both classes, I included a variety of questions so as not to focus attention entirely on this point.

Finally, for research question 3, the feasibility of incorporating thematic teaching in EFL programs, I conducted an interview with the partner teacher and documented my own observations.

Results

Written Homework

Analysis of the summaries revealed that the great majority of students in the non-TC class copied sentences directly from the story and linked them together in a way that was not conducive to smooth reading or understanding by a reader unfamiliar with the original. The TC class students, however, were more adept at creating an original topic sentence and explaining the story effectively, albeit with abundant grammatical errors.

Table 1 shows two examples that demonstrate this observation. The TC student who read this story in her English reading class was more adept at summarizing the story in her own words. She was able to use language in an independent and fluent manner while pushing her writing to the limits of her linguistic ability. However, the non-TC student relied on the structures of the original reading. Aside from changing pronouns, he did not exercise his ability to cognitively construct language.

Table 1

*Summary Paragraph of Impact Issue's Unit 10 Family Values***Sample student summary from the TC group:**

This sentence was written by Mi-Yeon when she was a little girl. That is when she was a little girl her mother's guest broken the vase. That vase is a precious vase. She was going and running that room. But her mother stopped her and her mother said, "It wasn't valuable." Her mother thinks, "We must not tell the truth" and "There is no other way."

Sample student summary from the Non-TC group:

She was in the kitchen helping her mother when she heard the crash. She knew at once what had happened. Our family vase has been broken. The vase that has been in our family for over 200 years is broken. We mustn't let our guests know how priceless it is. When we entered the room, she saw at once that my worst fear was right – the priceless vase was in pieces on the floor. And standing next to the broken vase was the 4 years old son of Mrs. Kim. She mother quickly said, "It was an old vase. It was not valuable."

Note. Passages taken directly from original story are underlined.

Review Test

Table 2 and Figure 1 show the test results and student answers to the questions in the *My Opinions* section. Analysis of the test scores revealed that the students in the TC class scored significantly higher ($p < .006$) than those in the non-TC class, with a mean of 75.9% ($N = 24$) and 64.9% ($N = 17$) respectively. The most marked difference can be seen in the final section, *My Opinions*, where students were required to write their opinion concerning six topics along with the reason for their stance. Answers were evaluated on how well the opinion and reasons for it were expressed. On average, students in the TC class expressed their opinions and reasons on about twice as many topics as those in the non-TC class.

Final Grades

A clear difference is seen in both grade point average and distribution of grades

between the two classes. Table 3 shows that students in the TC class had an average GPA of 3.2, significantly higher than non-TC students, who averaged 2.6. In the TC class, 83% of students received an A or a B for their final grade, while only 41% of students in the non-TC class received such grades.

Feedback on the Textbook

At the end of the course, a questionnaire was given to students in both classes with the following item: "I think *Impact Issues* should be used as the text for English 1 & 2 (reading) as well." Students ranked their level of agreement from 1 (low) to 5 (high). Although the distribution is almost identical between the two classes for responses of 2 through 5, the main point of difference was in the number of students who responded strongly against using the same textbook. Five of the students in the TC class marked 1 as compared to only 2 in the non-TC class (see Figure 3).

Roughly one third of the students who used *Impact Issues* in both classes responded negatively. Of those

Table 2

Final Test Scores

Group	Test score range						Mean
	40s	50s	60s	70s	80s	90s	
TC students ($N = 23$)	0	0	6	8	7	2	75.9
Non-TC students ($N = 16$)	1	1	6	7	1	0	64.9

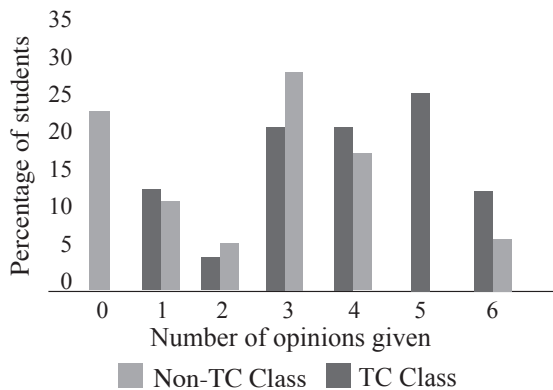


Figure 1. Students' expression of opinions

responding negatively, only one offered a reason—that she understood the content quickly. Quite possibly that student was more advanced than the others and became

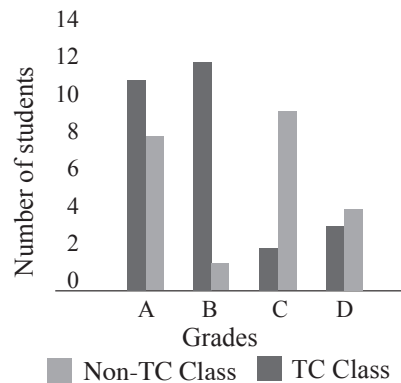


Figure 2. Grade distribution

were difficult; three suggested that Japanese translation should be included; two said the stories were too long; and one said the pace of the class was too fast.

Table 3
Final Grades

Class	Final grade (grade points)				Avg. GPA
	A (4)	B (3)	C (2)	D (1)	
TC (n = 28)	11	12	2	3	3.2
Non-TC (n = 22)	8	1	9	4	2.6

bored by the repetition of content. Nevertheless, 64% of the students' responses to using the same materials after experiencing one semester of such instruction were average or above. Seventy-eight percent of non-TC class students' responses were average or above.

The questionnaire also included space for written comments about *Impact Issues* (Questions 3 and 4). Out of 13 written comments from the TC group, only two were negative. Students responding negatively said that the topics were too serious. No students commented on the level of difficulty of the stories nor on the use of Japanese translation. In contrast, of 22 written comments from the non-TC group, 16 were negative: 10 students reported that the stories

When asked if they would prefer to use a textbook that included more conversational English rather than issues, the results were the opposite. The TC class (N = 22) was neutral with a 3.0 average and only 32% marking 4

or 5, yet the non-TC class (N = 25) had a much higher 4.1 average with 64%, marking 4 or 5.

Teacher Observations

Aside from statistical data, I recorded my own observations of differences between the two classes. The most poignant differences were overall student energy, enthusiasm, and pace of the activities.

One common technique used as a warm-up exercise was to ask short answer questions such as "Where was Mi-Yeon? Who was she with? What did she hear?", etc. I encouraged all students who thought they knew the answer to raise their hand, and I would wait until most students raised their hands before

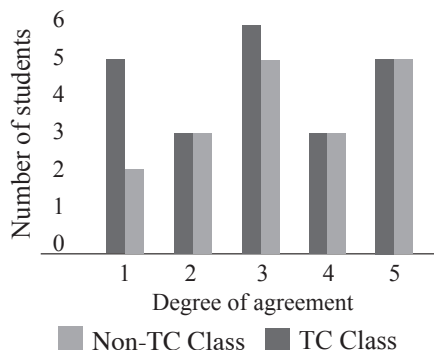


Figure 3. Item: I think *Impact Issues* should be used as the text for English 1 & 2 (reading) as well.

calling on one. In the TC class, students were able to raise hands almost instantaneously, and I was able to ask questions at a very rapid pace. The students in the non-TC class were two to three times slower in answering the questions and often had to resort to skimming the story again before raising their hands. The TC class students smiled and seemed to enjoy this activity, reacting as if it were a game, whereas the non-TC class students drudgely engaged in it as a tedious drill. Often, everyone was discouraged.

Discussion & Analysis

Student Performance

Given the significantly higher test results for the experimental group, the complementary use of the text materials may have a positive effect on students' overall ability to express themselves in English. It is of course logical that students who use the same materials in two 90-minute classes should score higher on a test based on that material than those who only used it half as long. However, indications that student ability to express their ideas in a fluent manner increased significantly from using the same materials in both classes can be deduced from the results from *My Opinions*.

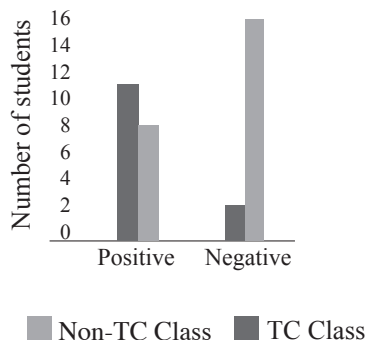


Figure 4. Student comments on using *Impact Issues*

That 74% of the TC class students scored in the 70s or above as compared to only 47% of the non-TC class suggests that thematic teaching may help students attain a deeper and more accurate understanding of the content, vocabulary and structures in the topical readings. Given the small sample size, though, it is not possible to come to a conclusive view that a thematic curriculum will necessarily lead to increased understanding of content and skills, but the results indicate that it may be beneficial for the students.

The recycling of language in multiple classes may contribute to students' increased language proficiency. One benefit is that language is automatically reviewed and used in a variety of ways. Research shows that students need to be exposed to vocabulary at least 15 to 20 times for it to be retained in long-term memory (Waring, 2002; Nation, 2001). In Rivers' (1975) framework for aiding students to transfer from initial acquaintance with language to mastery, the key is transferring old knowledge about well-known material to using new vocabulary and structures in new contexts. By allowing students to examine themes through tasks involving a variety of skills, they are thus given opportunities to use language in a variety of contexts. For instance, students in the TC group read the passages and prac-

ticed related reading skills building tasks in their Reading class. Then, they encountered the same vocabulary again when engaging in tasks in their Oral English class. Students need opportunities to re-use and re-structure memorized expressions, patterns and vocabulary and apply them in new contexts (Mitchell, 2003). Ideally, students will benefit from recycling topic-related language in their required language courses.

Another benefit of using the same materials in multiple classes is that students learn similar material under a variety of teacher styles and methods. Given the extensive research on individuality of learning styles (Gardner, 1993), it is not surprising that students can benefit from a variety of teaching styles while using the same base material. In the questionnaire, students commented that, "If there was something I didn't understand, I could understand at least one of the two teachers' explanations." And "It was good to experience two ways of using the same materials. I could have a deeper understanding of the materials." It would be interesting to discover how a thematic approach can appeal to various modalities of learning.

Student Preferences

Regarding questionnaire responses about using the same textbook in two classes, one can only imagine why so many students from the non-TC class wanted to use the same textbook. Quite possibly, they heard about the experiment from friends and wanted to be able to enjoy the same benefits (better understanding, less money spent on textbooks, etc). Further study may confirm this.

Considering the favorable responses of the TC class students versus the predominantly negative responses of the non-TC students towards the use of the textbook, one might conclude that the TC class, for the most

part, felt the topics were interesting and the textbook was valuable for English studies. However, the majority of the students who used the textbook only in their oral English class had difficulty focusing on both comprehension of the content and practicing oral communication skills. So a thematic curriculum might aid the oral communication teacher in conducting the class in English, and for students to maintain a high level of comprehension while communicating in English. Since students already had practice reading the stories and going over the meaning in Japanese in their reading class, they could focus on the process of expressing the content of the stories and their opinions in English. Using the same core materials in more than one class might be especially helpful for low level students.

The majority of non-TC students expressed that they would prefer a textbook with more conversation practice. As the teacher, one assumption I can make to account for this difference is that the TC group more easily engaged in conversation and discussion practice because they had more background knowledge of the content of the material. On the other hand, I had to take class time to explain the material before engaging in oral practice, thus considerably decreasing students' conversation practice time. This may explain why those students felt the need for a conversation based textbook.

Creation of a Thematic Curriculum

Finally, one goal of this study is to see if it is feasible to incorporate this practice in other university curriculums. The findings documented here are based on the reflections of the two teachers involved in this study.

In setting up this program, the main benefit for me as a teacher was the increased interaction with my colleague, leading to dis-

cussions of approaches to teaching. We met periodically to evaluate and discuss materials and the pace at which we would proceed. We also created a shared binder for materials (both main and supplementary), lesson plans and notes, homework, evaluation methods and results. In my experience, it is unfortunately rare to have such pedagogically based communication with a colleague.

There were of course several challenges. Choosing material that would be as equally conducive to teaching reading and structures as well as oral communication skills was difficult. At that time, many published topically based course books were either slanted towards practicing oral communication, reading or writing. No matter which book we chose, one teacher would inevitably be forced to create or search for more supplementary materials than the other. In our case, *Impact Issues* was chosen, though supplemental reading material was required. To ease this burden, we both took responsibility for creating a base of additional reading material. In terms of skills building practice, it was necessary to supplement in both areas.

Some researchers find it preferable for teachers to create materials tailor-fit to the students and even involve the students in the process (Jorstad, 1979). Some teachers may resent the added burden. This is why we followed the lead of Aoyama Gakuin University (Strong, 1995) and focused on published materials. We found the clear division of reading and speaking practice in *Impact Issues* and distinctive themes quite easy to work with. Another challenge was on an administrative level. We had to arrange the class schedule with the Academic Affairs Office to make sure that both of us taught the same group of students. So it is important to consider at what level a thematic curriculum is to be implemented. My experience suggests that

it would be easiest to implement at the program level. In that way, all instructors would be involved in the process—via meetings, email and memos—and would have access to materials organized in files in the teachers' lounge. Further research is needed to explore efficient and effective ways of implementing a thematic approach to curriculum design.

Conclusion

This study looked at two small classes over a one-year period. Thus the sample size is too small to draw broad conclusions. However, the generally positive results indicate that this practice may benefit both students and teachers. I hope that this study provides impetus for further research in this area.

To evaluate the effectiveness of this approach, longitudinal studies in a variety of contexts is necessary. For instance, studies in programs that require more than two courses in one academic year or programs for English majors could demonstrate how using thematic units aids curriculum design.

In addition, research related to collaboration between teachers might show how this approach helps or hinders relationships among colleagues. Initially, primarily commercially produced materials can be selected and supplemented with materials developed and gathered by teachers over time. By involving teachers in program development, communication among faculty members may be encouraged and a sense of investment in the program instilled.

In addition to using commercially published course books, teachers may select content books as a thematic base. This has become a common practice in learning communities across the curriculum in the US, such as the practice of balanced literacy (Johnson, 1999). Irujo (1990) points out that using content books as a base for a thematic

curriculum is one area that could be explored in foreign language education. As research indicates that extensive reading is an essential component of any language program (Krashen, 1993), using content books to establish topical themes across skills-based courses may be one way to encourage reading while recycling language and structures in a variety of contexts. This may be of particular interest to literature specialists.

Using a thematic approach to curriculum design over a long term can promote consistency across classes while accommodating the diverse needs and interests of students. In reference to bilingual immersion programs, Genesee (1994) suggests that content "... need not be academic; it can include any topic, theme, or non-language issue of interest or importance to the learners" (§ 8). How can we determine student interest, especially when teachers are in the initial planning stage of a thematic curriculum? This is another area that should be explored further.

I hope that this study plants seeds of thought for innovative ways that educators can adopt practices to suit the ever changing needs of students and institutions.

References

- Brown, J. D. (1995). *The elements of language curriculum: A systematic approach to program development*. Boston: Heinle & Heinle Publishers.
- Day, R., & Yamanaka, J. (1998). *Impact issues: Thirty key issues to help you express yourself in English*. Hong Kong: Longman.
- Day, R., & Yamanaka, J. (1999). *Impact topics: Thirty exciting topics to talk about in English*. Hong Kong: Longman.
- Galda, L. (1998). Mirrors and windows: Reading as transformation. In T.E. Raphael and K. H. Au (Eds.), *Literature-based instruction: Reshaping the curriculum* (pp. 1-11). Norwood, MA: Christorpher-Gordon.
- Gardner, H. (1993). *Multiple intelligences – The theory in practice*. New York: BasicBooks.
- Genesee, F. (1994). *Integrating language and content: Lessons from immersion*. Retrieved January 6, 2006, from Center for Research on Education, Diversity & Excellence. University of California, Santa Cruz. Web site: <http://repositories.cdlib.org/crede/ncrcdslleducational.EPR11>
- Goodman, Y.M. (1998). Roots of the whole-language movement. *Elementary School Journal*, 90(2), 113-127.
- Grisham, D. (1995). Exploring integrated curriculum. *Reading Psychology*, 16, 269-279.
- Hansen, J. (1997). Researchers in our own classrooms: What propels teacher researchers? In D. Leu, C. Kinzer, & K. Hinchman (Eds.) *Literacies for the 21st century: Research and practice* (pp. 1-14). Chicago: National Reading Conference.
- Irujo, S. (1990). *How to plan content-based teaching units for ESL*. Paper presented at the Annual Meeting of the Teachers of English to Speakers of Other Languages (TESOL, 24th, San Francisco, CA, March 6-10).
- Johnson, D. (1999). *Balanced reading instruction: Review of literature*. Retrieved January 6, 2006, from North Central Regional Educational Laboratory Web site: <http://www.ncrel.org/sdrs/timely/briiss.htm>
- Jorstad, H. (1979). Objectives for new programs: A thematic approach in second language learning. In J. Arendt et al. (Eds.) *Foreign language learning, today and tomorrow* (pp. 25-40). New York: Pergamon Press, Inc.
- Kagen, S. (1992). *Cooperative learning*. San Juan Capistrano, CA: Resources for Teachers.
- Kiyomi, S., Tsukamoto, M., Honda, A., & Sawazaki, R. (2004). An interim report on research in curriculum design for the

improvement of English education at Surugadai University. *Surugadai University Studies*, 28, 149-178.

Kouraogo, Pierre. (1987). Curriculum renewal and INSET in difficult circumstances. *ELT Journal*, 41(3), 171-178.

Krashen, S. (1993). *The power of reading: insights from the research*. Englewood, CO: Libraries Unlimited, Inc.

Mitchell, R. (2003). Rethinking the concept of progression in the national curriculum for modern foreign languages: A research perspective. *Language Learning Journal*, 27, 15-23.

Nation, P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.

Rivers, W. (1975). *A practical guide to the teaching of French*. New York: Oxford University Press.

Strong, G. (1995). Using a theme-based

curriculum to teach English. Unpublished teaching guide. (ERIC Document Reproduction Service No. ED 397 630)

Walmsley, S.A., & Walp, T.P. (1990).

Integrating literature and composing into the language arts curriculum: Philosophy and practice. *Elementary School Journal*, 90(3), 51-274.

Waring, R. (2002). Basic principles and practice in vocabulary instruction. *The Language Teacher*, 26(7), 11-13.

Wesche, M. B. (1993). Discipline-based approaches to language study: Research issues and outcomes. In M. Krueger & F. Ryan (Eds.) *Language and content: Discipline- and content-based approaches to language study* (pp. 57-82). Lexington, MA: D.C. Heath.

Note: I would like to express my appreciation to Professor Mieko Tsukamoto for her collaboration for this study.

Appendix

English 3 Feedback

Please rank the following from 1 (low) to 5 (high)

- _____ 1. I enjoyed this class.
- _____ 2. I understood what Renée said when she spoke in English.
- _____ 3. I wish Renée had spoken more English in class than she did.
- _____ 4. I can speak English better now than I could in April.
- _____ 5. I have more confidence in speaking English now than I did in April.
- _____ 6. I wish we had done more speeches and presentations.
- _____ 7. I think Impact Issues is useful for improving my English skills.
- _____ 8. I think Impact Issues should be used as the text for English 1 and English 2 as well.
- _____ 9. I would prefer to use a textbook that had more speaking practice.
- _____ 10. I would prefer to use a textbook that had more everyday English (conversational English) rather than topics/issues.

Please write any comments or suggestions:

1. What were the good/strong points of this spring semester?
2. What were the bad/weak points of this semester?
3. What were the good/strong points of our textbook, Impact Issues?
4. What were the bad/weak points of our textbook, Impact Issues?
5. What are you looking forward to doing in class in the fall semester?

Any other comments?

Research Digest

Quantitative Methods for Investigating Attitudes: The Matched-Guise Technique

Christopher Long
Tohoku Gakuin University

Introduction

As the recent anti-Japan demonstrations in China reveal, attitudes have a large impact on behavior. Anthropologists and sociologists have long recognized that attitudes constitute a fundamental aspect of any culture (e.g., Everett & Steinfatt, 1999; Ferraro, 2002). Moreover, the relationship between language and culture (e.g., attitudes) continues to spark interest and debate within the field of linguistics (e.g., Hall, 1976; Sapir, 1921).

This paper takes as its starting point the belief that attitudes are important indicators of behavioral intentions (for a discussion, see Everett & Steinfatt, 1999, p. 81) and that an understanding of attitudes can aid the study of human interaction across a broad range of disciplines (e.g., intercultural communication, linguistics, education).

The paper presents a brief overview of attitude studies focusing on an experimental method called the matched-guise technique (e.g., Lambert et al., 1960). The discussion is intended to be both theoretical and practical. It is theoretical in that it presents a critical overview of attitude studies while addressing core issues concerning the difficulties inherent in studying attitudes. It is practical in that the discussion is intended to give the reader the minimal knowledge necessary to

undertake a study of their own, using the techniques described.

Early Studies: Social Distance Scale

Attitudes are difficult to study because they are internal and thus not directly observable. According to Everett and Steinfatt (1999), attitudes are “emotional responses to objects, ideas, and people” (p. 81). Furthermore, although there is not necessarily a “one-to-one” relationship between attitudes and behavior, attitudes indicate “behavioral intentions” (Everett & Steinfatt, 1999, p. 81). Given the close relationship between attitudes and behavior, it is not surprising that most studies investigate verbal and physical behavior as an expression of attitudes (e.g., Bourhis & Giles, 1976; Bogardus, 1933; Choy & Dodd, 1976; Fielding & Evered, 1980; Lambert et al., 1960; Long et al., 2005).

In an early and influential study of attitudes, Bogardus (1933) developed a technique for investigating Park's (1924) concept of social distance. In Bogardus' study, participants were asked to respond to a series of questions designed to evaluate their perceived

intimacy/distance towards different ethnic groups. Questions included examples such as: “Would you marry someone who is Chinese?” “Would you have Chinese people as regular friends?” “Would you have Chinese people as speaking acquaintances?” (cited in Hudson, 1996, p. 212).

One criticism of Bogardus’ study is that he had participants give “overt” evaluations of different groups. In other words, they answered highly transparent questions regarding their attitudes. There are two problems with such responses. One is that respondents may not be aware of their true attitudes; attitudes can be illusive to both observer and holder. On the other hand, respondents may have concealed their true attitudes in their responses. This would seem particularly likely with responses that reveal negative attitudes. Therefore, although behavior can potentially reveal attitudes, not all behavior provides an equally valid measure. Particularly, the distinction between data derived from “overt” vs. “covert” experimental techniques appears to be significant.

An additional problem with Bogardus’ design is that participants responded to questions about “general” or “abstract” groups of people (i.e., Chinese people). As noted above, attitudes can be seen as “emotional responses” to objects, people, and ideas. Therefore, when assessing attitudes, it is vital that “responses” be as natural as possible. When responding to questions about general groups, participants do not provide first-hand reactions, but rather second-hand reports. Essentially they are being asked to summarize their past experiences. This is problematic, because it allows for editing, error, or both.

Further Developments: The Subjective Reaction Test

The subjective reaction test, developed by Lambert (1967) and Giles and Powesland (1975), overcomes these shortcomings of the Bogardus study. For the test, different individuals record an identical brief passage (e.g., a section from a text, a poem, a sentence) which participants then listen to and evaluate on number of scales. The scales consist of two contrasting adjectives (e.g., intelligent/unintelligent; trustworthy/untrustworthy, kind/unkind) placed at either end of a seven point continuum. Participants evaluate the speaker by selecting a point along the continuum.

Because participants evaluate specific aspects of individual speakers, the test avoids the problems inherent in asking for opinions of abstract groups. The test also allows for a quantitative assessment of attitudes (i.e., the numerical rating for each of the items).

Although participants rate individual voices, the ultimate goal of the technique is to compare participant evaluations of different groups. Therefore, the voices are clearly recognizable as belonging to particular social or ethnic groups (e.g., Black English, Cockney, New York working class). Although some studies employing this technique have considered differences in overall evaluation of speakers of standard vs. non-standard dialects (e.g., Paltridge & Giles, 1984), others have investigated evaluations within specific social contexts (e.g., Hopper & Williams, 1973; Kalin & Rayko, 1980; Labov, 1972). For example, Kalin and Rayko (1980) investigated Canadian university students’ impressions of the employability of accented and unaccented English speakers. They report that participants rated accented speakers as more suited for low-status positions (e.g., plant cleaner), and unaccented speakers as

more suited for high-status positions (e.g., foreman and industrial mechanic). Other studies have investigated teachers' evaluations of children who use standard vs. non-standard dialects (e.g., Seligman et al., 1972; Choy & Dodd, 1976). Choy and Dodd (1976) report that teachers consistently gave more favorable evaluations to students who spoke Standard English compared to students who used a Hawaiian dialect.

One criticism of the subjective reaction test is that it is difficult to determine whether differences in evaluation result from attitudes towards groups or attitudes towards the individual personalities of the voices used. For this reason, we must be cautious when generalizing the results of subjective reaction tests from the individual to group level.

Matched-Guise Technique

The matched-guise test maintains the basic format of the subjective reaction test but improves on it by making a simple modification: Rather than evaluating different individuals, participants listen to and evaluate the same individual in two different guises (e.g., when using two different speech varieties). Participants think that they are evaluating different individuals, and thus any differences in evaluation across guises can be safely assumed to result from attitudes towards the groups that the guises represent.

Matched-Guise Technique: Design Type 1

There are two basic designs for the matched-guise technique. In the first, the same speaker makes two recordings, each using a different language variety (e.g., French and English; Standard and non-standard English). The same group of participants then listens to and evaluates both guises,

thinking they are evaluating different individuals. Lambert et al. (1960) utilized this technique in their landmark investigation of language attitudes in Canada by having participants evaluate French and English guises. Similarly, Seggie et al. (1986) had an actor record an identical monologue in four different accents (guises) and then had groups evaluate the four "different" individuals. Using the same technique, Okamoto (2001) obtained evaluations of speakers of Standard Japanese and speakers of the Nagoya dialect of Japanese.

One advantage of this experimental design is that because the same group of participants evaluates both guises, any differences in evaluation can not be attributed to variation across participant groups. However, because participants evaluate different speech varieties (often times different languages), in some sense they are not actually evaluating the same individual. Thus, although the goal of the study is to assess differences in attitudes towards groups by using the same individual as a representative of different groups, differences in evaluation may in fact result from differences in voice, prosody, timing, pauses, and any one of a number of other factors that vary across language variety. In other words, like the subjective reaction test, this version of the matched-guise test is limited because it does not clearly distinguish between attitudes towards groups and attitudes towards individuals.

Matched-Guise Technique: Design Type 2

The second design type addresses the problem of distinguishing group attitudes from individual attitudes. In this design, the same individual makes a single recording which is then evaluated by two different

groups, each of whom receives different background information on the speaker (e.g., different nationality, ethnicity, etc.). Because both groups listen to and evaluate an identical passage, there is no possibility that differences in evaluation are the result of differences in language variety (e.g., voice, prosody, etc.). They can only be attributed to attitudes towards the groups to which the guises are presented as belonging.

Recent studies employing this technique have investigated Japanese attitudes towards foreigners. For example, Long et al. (2005) had two participant groups of Japanese university students listen to short self-introductions (in Japanese) by speakers they believed to be either a bilingual Chinese or bilingual American (recordings were made by Japanese native speakers). Results indicated that when participants believed they were listening to an American they rated the speaker more positively on 8 of the characteristics investigated (warm, cheerful, sense of humor, talkative, loud, easy to listen to, attractive, tall). "Chinese" voices were rated more positively for three characteristics (intelligent, calm, and thin). Sorensen and Maeda (2002) utilized a similar design to investigate Japanese impressions of intermediate level non-native speakers of Japanese. They had two participant groups listen to a list of voices for which half were presented as Asian American and the other half Caucasian. The guises were switched for each group. They report that Caucasian guises were evaluated more positively on a number of scales (e.g., humorous, sociable, intelligent, imaginative, talkative, attractive, louder, and faster).

One problem with this second version of the matched-guise technique is that because separate groups listen to and evaluate each of the two guises, differences in evaluation may

be the result of base differences in the two groups. There are two ways to counter this potential problem. The first is to include a speaker whose guise remains constant across both groups. For example, in addition to 12 voices (6 Caucasian and 6 Asian American guises) that they switched for two groups of listeners, Sorensen and Maeda (2002) included three voices that remained constant across groups (one Caucasian, one Black, and one Asian). When analyzing the data, because no significant group differences were found for the guises that remained constant, they concluded that the groups could be safely compared.

Another way to control for group variation is to combine data across groups before running any analyses. For example, Long et al. (2005) played tapes of two voices, one in each guise (American and Chinese) and switched the guises for each group. They then combined the American and Chinese data for both groups before running their analyses. In this way, any systematic differences across the two groups are cancelled out.

Summary

In this paper I have argued that attitudes are important indicators of behavior and thus warrant investigation. I have presented a brief overview of attitude studies while considering the problems inherent in investigating attitudes. The main goal of this paper has been to introduce the matched-guise technique and consider the strengths and weaknesses of its two main design types. It is my hope that this discussion will have encouraged scholars to incorporate attitude studies into their own research. Given the current political climate in Japan and the world, it is particularly important to deepen our understanding of attitudes across a variety of cultures.

References

- Bogardus, E. S. (1933). A social distance scale. *Sociology and Social Research*, 17, 265-71.
- Bourhis, R. Y., & Giles, H. (1976). The language of co-operation in Wales: A field study. *Language Sciences*, 42, 13-16.
- Choy, S., & Dodd, D. (1976). Standard-English speaking and non-standard Hawaiian-English speaking children: Comprehension of both dialects and teachers' evaluations. *Journal of Educational Psychology*, 68, 184-193.
- Everett, M. R., & Steinfatt, T. M., (1999). *Intercultural communication*. Prospect Heights, Illinois: Waveland Press.
- Ferraro, G. P. (2002). *The cultural dimension of intercultural business* (4th ed.). New Jersey: Pearson Education.
- Fielding, G., & Evered, C. (1980). The influence of patients' speech upon doctors. In R. N. St. Clair & H. Giles (Eds.), *The Social and psychological contexts of language* (pp. 51-72). Hillsdale, NJ: Earlbaum.
- Giles, H., & Powesland, P. (1975). *Speech style and social evaluation*. London: Academic Press.
- Hall, E. (1976). *Beyond culture*. New York: Random House.
- Hopper, R., & Williams, F. (1973). Speech characteristics and employability. *Speech Monographs*, 46, 296-302.
- Hudson, R. A. (1996). *Sociolinguistics* (2nd ed.). London: Cambridge University Press.
- Kalin, R., & Rayko, D. (1980). The social significance of speech in the job interview. In R. N. St Clair & H. Giles (Eds.), *The social and psychological contexts of language* (pp. 39-50). Hillside, NJ: Earlbaum.
- Labov, W. (1972). *Sociolinguistic patterns*. Oxford: Blackwell.
- Lambert, W. E. (1967). A social psychology of bilingualism. *Journal of Social Issues*, 23, 91-109.
- Lambert, W. E., Hodgson, R. C., Gardner, R. C., & Fillenbaum, S. (1960). Evaluational reactions to spoken languages. *Journal of Abnormal and Social Psychology*, 60(1), 44-51.
- Long, C., Otsuka, H., & Ikeda, N. (2005, January). *Japanese university students' stereotypes of Chinese and American foreigners: A quantitative study*. Paper presented at the meeting of the Japan Association for Language Teaching, Yokohama Chapter, Japan.
- Okamoto, S. (2001). Nagoya hougen no shiyou ga hanashite no inshou ni oyobosu eikyou [Effects of Nagoya Dialect on the Impressions of Speakers: Using the Matched-guise Technique]. *The Japanese Journal of Language in Society*, 3(2), 4-16.
- Paltridge, J. & Giles, H. (1984). Attitudes towards speakers of regional accents of French: Effects of regionality, age, and sex of listeners. *Linguistische Berichte*, 90, 71-85.
- Park, R. E. (1924). The concept of social distance. *Journal of Applied Sociology*, 8, 339-44.
- Sapir, E. (1921). *Language: An introduction to the study of speech*. New York: Harcourt & Brace.
- Seggie, I., Smith, N., & Hodgins, P. (1986). Evaluations of employment suitability based on accent alone: An Australian case study. *Language Sciences*, 8, 129-140.
- Seligman, C., Tucker, G., & Lambert, W. (1972). The effects of speech style and other attributes on teachers' attitudes towards pupils. *Language in Society*, 1, 131-142.
- Sorensen C., & Maeda, N. (2002). Providing a classroom experience to learn about stereotypes: How ethnic appearance effects native speaker attitudes towards non-native speakers of Japanese. *Kanda Gaigo Daigaku Kiyou [Journal of Kanda University of Foreign Languages]*, 14, 409-437.

Use of Metatext in English and Japanese Research Articles: A Contrastive Study

Kaoru Kobayashi

Tokyo University of Agriculture

Tokyo University of Pharmacy and Life Science

Tamagawa University

Introduction

A key to success in academic writing is to maintain coherence of the text. One of the coherence devices is metatext, which guides the readers throughout the logical development of the text. It has been clarified that Finnish speakers and Spanish speakers use less metatext when writing academic papers than English speakers (Mauranen, 1993; Valero-Garcés, 1996). This paper reports on a statistical analysis of differences in the use of metatext between article sections of English and Japanese research in the field of applied linguistics; specifically, the *abstract*, *introduction*, *background*, *results and discussion*, and *conclusion* sections. The purpose is to see whether the logical progression of text is handled differently between the two languages. This can be the first step to projects that aim at developing better ways of teaching English academic writing to Japanese speakers.

Basic Concept

Mauranen (1993) defines metatext as “essentially text about the text itself ... (that) serves to organize the propositional content of the text and to comment on it” (pp.7-8). More recently, Hyland (2000), using the term *metadiscourse* to refer to metatext, defines it as “those aspects of the text which explicitly

refer to the organization of the discourse or the writer’s stance towards either its content or the reader” (p. 109). Metatext, in other words, is text that does not affect the propositional contents, but indicates the relationships between them.

The study of metatext derives from text linguistics in which coherence is the main focus of research. Recently, metatext has been studied within the realm of contrastive rhetoric because the use of metatext has been found to be cultural. Crismore, Markkanen and Steffensen (1993) claim that the study of metatext (metadiscourse, in their text) between different languages is meaningful because it may reveal its cross-linguistic features, linguistic and cultural differences, and better understanding of metatext as a rhetorical device. Metatext study, therefore, can be pedagogically important because it will help learners to be more conscious of the textual differences in organization, and the coherence differences between their native language and the target language.

Most studies on metatext between two languages involve English, and find that the use of metatext varies with language. In 1993, Mauranen studied the use of metatext in Finnish and English economics texts, and discussed how English-speaking scholars used more metatext than Finnish speaking scholars. Finnish text showed more *reader-*

responsible features, in other words, textual reliance on assumed shared knowledge that is placed on the reader. Comparing the use of metatext in economics texts written in English, between Anglo-American writers and Spanish-speaking writers, Valero-Garcés (1996) also claimed that Anglo-American writers were more careful about orienting the reader by using more metatext.

Methods

Data

A corpus of 60 applied linguistics experimental research articles (30 in English and 30 in Japanese) that closely conform to the conventional *Introduction-Methods-Results-Discussion* (I-M-R-D) pattern was selected from journals in the field of language teaching and learning, published in the years 1995 to 2000. Four journals were selected for their popularity among researchers in the respective fields. They were *TESOL Quarterly*, *Studies in Second Language Acquisition*, *Nihongo Kyoiku* and *Nihongo Kyoiku Ronshu*. Among the 30 articles written in English, 15 were obtained from *TESOL Quarterly* and the other 15 were obtained from *Studies in Second Language Acquisition*. Whether the writers are native speakers of English or not is unknown, but the intended audience are fluent readers of English.

Among the 30 Japanese articles, 26 were obtained from *Nihongo Kyoiku* (Journal of Japanese Language Teaching), and the remaining four were obtained from *Nihongo Kyoiku Ronshu* (Journal of Japanese Language Teaching). The titles of these two journals are incidentally similar, but they are published by different organizations. *Nihongo Kyoiku* is published by the Society for Teaching Japanese as a Foreign Language. *Nihongo Kyoiku Ronshu* is published by the

University of Tsukuba. Among these articles, two were written by authors whose names are not Japanese. Whether these writers are native speakers of Japanese or not is unknown, but their intended audience are fluent readers of Japanese.

Procedures

First, the structure of each article was analyzed. Among the *abstract, introduction, background, methods, results and discussion*, and *conclusion* sections, 23 English articles and 13 Japanese articles had the *introduction* section, and 26 English and 27 Japanese articles had the *conclusion* section. The *background, methods, results and discussion* sections were found in all the articles studied (Kobayashi, 2005).

Second, for this study, Mauranen's definition of metatext, which appears to be the more concrete among those presented by major researchers including Vande Kopple (1985), Nash (1992), Crismore, Markkanen and Steffensen (1993), was adopted. Mauranen categorized metatext into four kinds as follows:

Connectors: Conjunctions, adverbial and prepositional phrases which indicate the relationships between propositions in the text: *however, for example, as a result, ...*

Reviews: Clauses (sometimes abbreviated), which contain an explicit indicator that an earlier stage of the text is being repeated or summarized: *So far we have assumed that the corporate tax is a proportional tax on economic income.*

Previews: Clauses (sometimes abbreviated), which contain an explicit indicator that a later stage of the text is being anticipated: *We show below that each of the initial owners will find this policy*

to be utility maximizing.

Action Markers: Indicators of discourse acts performed in the text: *the explanation is, to express this argument in notation, to illustrate the size of this distortion,...*

(Mauranen, 1993, pp. 9-10)

Since this study focused on the logical progression of text, action markers were not included because action markers—unlike connectors, reviews and previews—do not directly show how one semantic segment relates to another, but rather indicate what the author is trying to express when presenting a particular semantic segment.

Each structural section of the articles was divided into semantic segments. A semantic segment in this study is similar to a *move* or a *step* in Swales' (1990) studies. The following examples show *connectors*, *reviews*, and *previews* from the articles.

Connector

Connectors indicate the kind of semantic linkage between the semantic segments.

- (1) The question, *then*, is whether it is justifiable to include stalling strategies among CSs or not.

In this example, “then” indicates a causal relationship between the prior semantic segment and the following one.

Review

A review signifies that the semantic segment following it refers to the prior segment.

- (2) *To summarize*, both providing background knowledge and previewing have been shown to be effective for both L1 and L2 readers.

“To summarize” is a review that announces the beginning of the semantic segment which is a summary of the prior semantic

segment.

Preview

A preview guides the reader to the following textual stage.

- (3) *A brief summary of the problem and the arguments follows.*

“A brief summary of the problem and the arguments follows” is a preview that foresees that the following semantic segment is a summary of the previous text.

In the third step of this study, the density of each category of metatext was analyzed by calculating the number of metatextual phrases per semantic segment. Then, to gain a better understanding of the difference between the density of each category of metatext between the two sets of data, the non-parametric Mann-Whitney U test was used because no specific type of distribution was assumed across the population variance. A *p* value less than 0.05 denies the following null hypothesis:

- $$H_0: \text{The density of metatext between the English and Japanese articles have the same mean and variance.}$$

Results

The results are given in Tables 1, 2, and 3. The calculation of *p* involves mean and data variance, so the greater difference in mean does not necessarily indicate that they are from different populations with different means and variances. Whereas “Yes” under “Difference” means that English and Japanese articles have different density, mean, and variance, “No” means that the two sets of articles share the same characteristics across the population.

Statistical differences were observed in the density of connectors in the *abstract* section, and the density of previews in the *abstract* and the *results and discussion* sections.

Table 1
The Density of “connectors” in
the English and Japanese Data

Section		Mean	z	p	Difference
Abstract	English	0.11	-4.73	.0001	Yes
	Japanese	0.41			
Introduction	English	0.15	-0.64	.5254	No
	Japanese	0.20			
Background	English	0.16	-1.56	.1179	No
	Japanese	0.26			
Method	English	0.05	-0.11	.9097	No
	Japanese	0.08			
Results and discussion	English	0.20	-1.93	.0536	No
	Japanese	0.27			
Conclusion	English	0.32	-1.36	.1742	No
	Japanese	0.27			

Table 2
The Density of “reviews” in
the English and Japanese Data

Section		Mean	z	p	Difference
Abstract	English	0.00	—	—	No
	Japanese	0.00			
Introduction	English	0.00	—	—	No
	Japanese	0.00			
Background	English	0.03	-1.63	.1026	No
	Japanese	0.01			
Method	English	0.01	-1.01	.3132	No
	Japanese	0.00			
Results and discussion	English	0.01	-0.08	.9394	No
	Japanese	0.01			
Conclusion	English	0.01	-1.30	.1935	No
	Japanese	0.05			

Based on mean values shown in Tables 1 and 3, the differences indicate a higher density of metatext in these article sections written in Japanese.

Discussion

The extensive use of metatext among the *abstract* sections written in Japanese may indicate that the format of the section is not established. Bhatia (1993) claims that the

abstract section in an English research article consists of four moves: Move 1, introducing purpose, Move 2, describing methodology, Move 3, summarizing results, Move 4, presenting conclusions. For the *abstract* section in Japanese research articles, there has not been a major study that describes the move structure.

Samples 1 and 2, taken from the research articles analyzed in this study, show the difference between English and Japanese approach to orienting the reader through the logical flow in the *abstract*. Move number on the right indicates the move in Bhatia's model that corresponds to the segment. Italicized phrases in the examples indicate metatext.

Sample 1

The *abstract* in Sample 1 can be divided into four semantic segments which

are in accordance with Moves 1 to 3 in Bhatia's model. The first segment introduces the study, the second and third segments explain the procedure, and the last segment presents the results. These four segments are not linked with metatext. Here, in order to guide the reader through the flow of the text, the writer resorts to the established generic structure of the *abstract* which is shared by both the writer and reader.

Table 3
The Density of “previews” in the English and Japanese Data

Section		Mean	<i>z</i>	<i>p</i>	Difference
Abstract	English	0.00	-2.31	.0207	Yes
	Japanese	0.02			
Introduction	English	0.00	-1.91	.0564	No
	Japanese	0.04			
Background	English	0.05	-0.28	.7829	No
	Japanese	0.06			
Method	English	0.01	-1.17	.2434	No
	Japanese	0.02			
Results and discussion	English	0.01	-2.09	.0367	Yes
	Japanese	0.03			
Conclusion	English	0.07	-1.93	.0534	No
	Japanese	0.01			

Sample 2

The five segments, indicated within Sample 2 are slightly different from Bhatia’s model. It starts with the introduction of the subject field, which is not found in Bhatia’s model. The second sentence begins as Move 1 but develops into Move 2, involving the third and fourth sentences. The rest of the section can be divided into Moves 3 and 4. Despite the fact that the number of words is usually limited for abstracts, the Japanese writer used metatext to link the segments between the second and third segments, and the third and

fourth segments. Perhaps the writer intended meta-text to guide the reader through the flow of logic because there is no set format for abstracts in Japanese writing. In fact, in a moves analysis study of abstracts using the same articles, a larger number of English articles showed the same pattern in the moves compared with the Japanese articles (Kobayashi, 2005).

The results show the use of metatext was more

frequently observed in the Japanese *results and discussion* sections than in the English *results and discussion* sections. In fact, more set patterns in the moves have been found in the English *results and discussion* sections (Kobayashi, 2005), which could indicate that the format of the section is more established among English speakers. It can be concluded that there is a tendency among the writers of Japanese texts to use metatext to maintain the logical progression of text.

Sample 1

English Abstract

This study replicates VanPatten and Cadierno (1993) in an attempt to determine whether or not explicit information given to learners receiving processing instruction is responsible for the beneficial effects of instruction.	Move 1
Fifty-nine subjects were divided into three groups: (1) one receiving processing instruction in object pronouns and word order in Spanish as in VanPatten and Cadierno (1993), (2) another receiving explanation only with no activities or practice, (3) and another receiving only the structured input activities with no explanation.	Move 2
A pretest/post-test assessment was used involving two tests, an interpretation test and a sentence-level production test.	Move 2
Results showed that the beneficial effects of instruction were due to the structured input activities and not to the explicit	Move 3

Sample 2

Japanese *Abstract* (and translation)

第二言語によるコミュニケーション能力を説明する理論的枠組みには、言語知識そのものと、それを操作して発話する方略的能力がある。本研究では、方略的能力に注目し、その方略的能力にかかわるコミュニケーション方略のトレーニングを行い、方略的能力の変容と言語知識との関連性を分析した。方略的能力は、コミュニケーション方略を用いて発話上生じた問題を解決した伝達性効率で測定した。言語知識はトレーニング前、後の語彙、文法、読解、聴解テストにより測定した。その結果、(1)トレーニング前半期から後半期における方略的能力の伸びに有意差が見られた。(2)トレーニング前の方略的能力と言語知識との相関は高かったが、トレーニング後には低くなることわかった。以上から、コミュニケーション方略のトレーニングにより方略的能力が向上すれば、限られた言語知識であっても、これを有効に活用することができると予想される。

Second language communicative competence is determined by both linguistic knowledge and strategic competence. In this paper we focus on strategic competence and provide (the learners) with a training course in communicative strategy, and then analyze the resulting changes in strategic competence and their relationship with linguistic knowledge. The strategic competence was measured by effectiveness of the communication strategy used to solve the communication problems. The language competence was measured by pre- and post-tests on words, grammar, reading and listening.

As a result, it was found that: (1) A comparison between the first and second four-week periods showed a significant improvement in strategic competence, (2) Strategic competence measured during the first period correlated well with linguistic knowledge measured in the pre-test, while strategic competence measured during the second period did not correlate well with linguistic knowledge measured in the post-test.

From the above, we conclude that even when a learner's linguistic knowledge is low, training in strategic competence will result in more effective communication.

Note. Translation by the author.

Conclusion

The results showed more metatext in the Japanese articles than in the English articles. This could indicate that Japanese speaking writers have used metatext to compensate for the lack of an established format of the *abstract* and *results and discussion* sections. This study has provided evidence for the difference in the way Japanese speaking writers and writers of English texts orient the readers through the textual progression. Future research can be done on research articles written in English by Japanese speakers to further clarify the use of metatext by Japanese speakers. More cross cultural studies of the use of metatext, with an analysis on

generic structure of the text, is suggested for a better teaching of English academic writing to Japanese speaking students.

References

- Bhatia, V. (1993). *Analysing genre: Language use in professionals settings*. London: Longman.
- Crismore, A., Markkanen, R., & Steffensen, M. S. (1993). Metadiscourse in persuasive writing. *Written Communication*, 10, 39-71.
- Hyland, K. (2000). *Disciplinary discourses: Social interactions in academic writing*. Essex: Pearson Education Limited.

- Kobayashi, K. (2005). The Structure of English and Japanese Research Articles: A Contrastive Study. *On CUE*, 13(2), 26-32.
- Mauranen, A. (1993). Contrastive ESP rhetoric: Metatext in Finnish-English economics texts. *English for Specific Purposes*, 12, 3-22.
- Nash, W. (1992). *An uncommon tongue*. London: Routledge.
- Swales, J. (1990). *Genre analysis: English in academic and research settings*. Cambridge: Cambridge University Press.
- Valero-Garcés, C. (1996). Contrastive ESP rhetoric: Metatext in Spanish-English economics texts. *English for Specific Purposes*, 15, 279-294.
- Vande Kopple, W. J. (1985). Some exploratory discourse on metadiscourse. *College Composition and Communication*, 36, 82-93.

Opinion & Perspective

The Teacher as Missionary: A Rebuttal

David Peaty
Ritsumeikan University

Readers of *The Language Teacher* may recall various articles in the last few years focusing on the issue of advocacy in global education (Guest, 2005; Peaty, 2004, 2005; Sargent, 2004). The article by Sargent (2004) criticized certain opinions, including mine, on this issue, and my response to Sargent (Peaty, 2004) clarified my views. This was followed by a hostile article from Guest (2005a), which I rebutted (Peaty, 2005). Guest has since repeated many of his accusations in a longer paper (2005b), and it is to this that I now respond.

In his latest article, Guest inadvertently demonstrates the need for critical thinking education by providing us with examples of fallacy, exaggeration, misrepresentation and irrelevance. His main accusation appears to be that I seek to promote a 'missionary' approach to classroom instruction.

In my own article (2004), I argued that there are limited situations in which teachers are justified in promoting certain viewpoints, values and principles. My first example was the Universal Declaration of Human Rights (UDHR), an international convention, the preamble of which urges teachers to "promote respect for these rights and freedoms." Guest evidently has no counter argument to this, and instead attempts to circumvent my argument by making the bizarre claim that because I think my political views coincide

with the goals of the UDHR, I assume I have a mandate to teach them. What I actually wrote, however, had nothing to do with either my own political views or my own teaching. All teachers, regardless of their political views, have a mandate to teach about the rights embodied in the UDHR. That is by no means the only example of an international convention which provides teachers with a mandate to promote a specific issue: others endorse education for peace, conservation, sustainable development and the elimination of racism and discrimination, to cite just a few examples.

Far from promoting a 'missionary approach' in the classroom, I merely state in my article that "there are contexts in which absolute neutrality is neither possible nor desirable" (p. 17). A similar observation can be found in Crick (1998, p. 56): "Of course, educators must never set out to indoctrinate, but to be completely unbiased is simply not possible, and on some issues, such as those concerning human rights, it is not desirable." I maintain that there are *certain* values and viewpoints which a teacher may - indeed should - present as *right*, without hedging or offering opposing views. Would Guest see fit to balance a condemnation of slavery or bonded labour, for example, with a defense of the practice? We may, of course, disagree on which values and viewpoints are beyond

dispute. However, this should not intimidate us into sitting on every fence.

I should perhaps remind the reader that the Crick Report was commissioned by the British government, and had a major influence on the citizenship curriculum introduced soon afterwards. I must also add that, contrary to Guest's claim, Crick does not inform us that "teaching only one side of an issue is...illegal in Britain" on page 59 or any other page. However, the Education Act 1996 does require school governing bodies, head teachers and local education authorities to "ensure that, where political or controversial issues are brought to pupils' attention, they are offered a balanced presentation or opposing views" (Qualifications and Curriculum Authority, 2000, p. 35). On the other hand, numerous exceptions become evident from even a cursory reading of the National Curriculum and related documents. In an official booklet advising teachers in England on citizenship education, for example, the QCA writes: "The need for balance should not be regarded as inhibiting a clear stance against racism and other forms of discrimination" (p. 35).

While Crick provides valuable insights into the issue of advocacy, we should bear in mind that both the Crick Report and the British education laws refer to compulsory education up to the age of 16. Sargent, Guest and I have so far neglected to identify the age group to which our arguments relate. However, my own experience is limited to university classes, and I would recognize that there is a much greater need for caution when presenting controversial issues to younger, more vulnerable learners.

Returning now to Guest's criticisms, I note with surprise his statement that "Peaty further justifies his use of the classroom as a podium for espousing this type of personal

dogma" (pp. 38-39). My article provides no indication whatsoever of who, what or how I teach. Guest is jumping to conclusions, and his conclusions are wrong. I have already stated categorically that "I do not believe in imposing my own views on my students" (2005, p.13), and Guest is in no position to dispute this.

Guest makes numerous references to my political views. For example, "Peaty goes on to characterize his views as non-mainstream, alternative, even radical" (p. 38) and "He seeks 'balance' by bringing in his 'marginal' or 'radical' perspectives" (p. 39). Guest later goes on to claim that these views "including all those advocated by Peaty" (p.39) can easily be found in the mainstream media. However, my article says nothing about my own political views. When I refer to reminders by eminent scientists that "our present way of life is unsustainable" (p. 16), I am making a statement of scientific fact, not of political opinion. I support this statement by referring to the United Nations Environment Program, the Intergovernmental Panel on Climate Change and the Worldwatch Institute, although Guest apparently does not recognize these institutions as authoritative. For his benefit, I will add a more specific citation: the Millennium Ecosystem Assessment Report, issued 31 March, 2005, by the Millennium Assessment Board on behalf of the United Nations, which warned that 15 of 24 global ecosystems are in decline. (The UN study is a synthesis of the work of about 1,300 researchers from 95 countries.) When referring to the "effects of consumerism on the global environment" (p. 16), I express neither approval nor disapproval. When I refer to Postman and Weingartner (1969) and Brown (1994), commenting that "each saw a need for the dominant or mainstream perspective to be challenged because busi-

ness as usual was clearly unsustainable” (p. 16), I am presenting their views, not mine. Even when I mention fair trade, myths about hunger and provocative questions about government policies (p. 17), I am not making any political statements. In short, there is nothing in my article that defines my own political views; Guest is again jumping to conclusions.

I quote from Stradling (1989, p.99) and Anderson (1996, p.24) in suggesting that there are perspectives that are avoided by the mainstream media. Guest inflates this into a wild accusation that I am ‘dispersing’ myths about a “singular, united, monolithic media” (p.39). In imposing his own concept of ‘progressive views’ on this discussion and then claiming that the ‘mainstream’ media are in fact progressive, or that supposedly radical issues are now mainstream, Guest appears to be arguing with himself over a red herring. According to him, the Daily Yomiuri is ‘right-wing’, ‘mainstream’ and ‘progressive’ - a curious blend indeed. He offers support only for the ‘progressive’ element: isolated articles about farm subsidies and whale habitats. However, two stories reproduced from other newspapers over a period of two days do not define the political position of a newspaper, as Guest should know; nor do they amount to evidence that there are no positions that the mainstream media avoid. I have subscribed to the Daily Yomiuri for more than 20 years, but can recall only one report dealing with the concept of fair trade (the example which I mentioned in my original article), and I have yet to meet a student who had even heard of this concept before taking my class. As to why the Daily Yomiuri has not offered greater exposure to this topic, I would assume it is due to lack of awareness or indifference, rather than any deliberate policy.

Guest tells us that opening the door to radical or marginalized groups would allow access to the views of Aryan supremacists (p. 39). This is an old, worn-out fallacy. Just because a school policy allows teachers to present views that are supported by scientific evidence and that threaten nobody that does not mean the school also has to permit the promotion of unscientific views which provoke violence against vulnerable communities. It is irrational to equate the message that consumerism harms the environment with the message of Aryan supremacists and unrepentant Stalinists. Moreover, Guest seems to have very little faith in the integrity of policy makers and supervisors. In fact, ratification of the United Nations International Convention on All Forms of Racial Discrimination should ensure that no school allows the promotion of Aryan supremacist views.

Rather than exhaust the reader with additional examples of exaggerations, misrepresentations and defective logic in Guest’s paper, I will turn now to the few positive contributions made in his article. The first is his observation that learners think independently and have well-founded opinions (p. 40). I have noticed the same thing, and therefore question the need for teachers to always coddle them with ‘balanced’ views. My university students can generally judge for themselves whether or not the resources I have chosen to use in class advocate a certain position and whether or not they accept it. However, there are gaps in their knowledge and awareness, as there are gaps in mine. These gaps are especially noticeable in relation to global issues, possibly because relatively few Japanese students read newspapers. Until I tell them, my students don’t know about the Universal Declaration of Human Rights or the Ramsar Convention. They have never heard of micro-credit or

the fair trade movement. Providing them with information about international treaties and innovative approaches to improving the lives of poor people is not indoctrination, as Guest would have us believe; it is global education.

This leads me to Guest's second contribution: the suggestion that students be invited to research 'topics'. This is hardly innovative: like many other teachers, and for many years, I have had students do research on various topics and present or discuss them in class or in research papers. Moreover, while they are talking, I take notes and learn from them, surely something not normally associated with the 'missionary approach.' However, I also feel a need, as a language teacher, to engage students in reading and listening tasks in the classroom in order to promote the development of language processing strategies and academic skills such as reading for main ideas, listening for gist, note taking and analysis of content. This requires resources such as magazine articles and television documentaries, and also short talks by the teacher. When such resources deal with global issues, they are seldom absolutely neutral. I believe that teachers should generally select resources that promote better global citizenship, and not attempt to balance them, as Guest would apparently require, with articles that encourage attitudes and behaviour that are detrimental to the environment, to social harmony and to global peace.

While Guest has greatly exaggerated the issue of advocacy, there are certain risks more worthy of our attention. Although it is acceptable to discuss the harmful effects of automobiles on the global environment and consider less harmful alternatives, discussion of the ethical failures of a certain car manufacturer that knowingly sold dangerous

products could be traumatic for a student whose father or mother worked for that company. Before we discuss the issue of social discrimination or sexual harassment in our classes, we need to consider the possibility that among our students there may be both victims and perpetrators. Discussion of such issues requires a higher level of sensitivity on the part of the teacher to signals from individual students suggesting the topic is causing distress.

As mentioned in my original article, it is also essential for teachers to constantly verify and update their information. They should also create a classroom environment in which students feel free to challenge the facts, assertions and opinions presented by teachers and class resources, and in which teachers feel no shame in admitting mistakes. Finally, before tackling risky subjects, teachers should confer with colleagues and coordinators to ensure that they are not violating institutional guidelines or commonsense parameters. Provided these basic precautions are taken, I believe that teachers at all levels are, and should remain, free to introduce ideas and viewpoints conducive to making the world a better place.

References

- Anderson, G. (1996). Global issues in the university ESL classroom. *The Language Teacher*, 20(11), 20-25.
- Brown, H.D. (1994). *Teaching global interdependence as a subversive activity*. Paper presented at the Georgetown University Round Table on Languages and Linguistics, Washington D.C., March 14th, 1994.
- Crick, B. (1998). *Education for citizenship and the teaching of democracy in schools. Final report of the Advisory Group on Citizenship*. London: Qualifications and Curriculum Authority.

- Guest, M. (2005a). A response to Peaty's "Global issues in EFL: Education or indoctrination?" *The Language Teacher*, 29(5), 11-12.
- Guest, M. (2005b). The teacher as missionary: Is this really global education? *On CUE*, 13(2), 38-42.
- Millennium Assessment Board. (2005). The Millennium Ecosystem Assessment: Statement of the MA Board. Retrieved on November 20th, 2005, from <<http://www.millenniumassessment.org/>>.
- Peaty, D. (2004). Global issues in EFL: Education or indoctrination? *The Language Teacher*, 28(8), 15-18.
- Peaty, D. (2005). A rebuttal to Michael Guest. *The Language Teacher*, 29(5), 12-13.
- Postman, N. & Weingartner, C. (1969). *Teaching as a subversive activity*. NY: Dell Publishing Company.
- Qualifications and Curriculum Authority (2000). *Citizenship at key stages 3 and 4: Initial guidance for schools*. London: Qualifications and Curriculum Authority.
- Sargent, T. (2004). Advocacy-oriented global education in TLT special issues. *The Language Teacher*, 28(2), 9-13.
- Stradling, B. (1989). How's your crap detector? In D. Randle (Ed.), *Teaching Green: A guide to education for life on earth* (pp. 98-99). London: Green Print.

Conference Review

A Tale of Two Delegates: JALT National Conference 2005: Sharing Our Stories

*David Ockert
Shiga University*

*Jerry Talandis Jr.
Toyama College of Foreign Languages*

Held at the Granship Convention & Arts Centre in Shizuoka, Japan from October 7 – 10, 2005, this year's JALT National Conference more than met participants' and presenters' expectations. We arrived on Friday evening just in time to meet JALT's President Steve Brown on an early retreat from the evening's festivities, since he had dinner arrangements with the plenary speakers. We still managed to find a tasty beverage at the bar and enjoyed chatting with both friends old and new in the relaxing, dimly-lit atmosphere.

The Tenji Hall was crowded on both Saturday and Sunday from 1 to 3 pm with poster presenters answering questions of visitors when they themselves were not wandering back and forth between all of the colorful presentations. On Saturday, Kim Bradford-Watts was busy answering "100 Questions to Ask before you Publish a Textbook." Also, Steve Martin and his colleagues had several inquiries about their presentation on "Motivation and the Event-Driven Curriculum."

Jerry attended "Teachers Who Moved Us" by Tim Murphey, and Dave Ragan and

he writes: I went to this presentation pretty much because of Tim Murphey. I had attended a presentation of his last year in Nara, and really loved it. This time, even though the topic was not on my 'to do' list, I felt that 'going with the good presenter' was a smart approach, and I wasn't disappointed. The workshop was about inspirational teachers and the reason they were inspirational. The presenters talked about the importance of stories and led us in a discussion about teachers who had made a difference in our lives. Murphey noted that stories are "healthy, safe, and inspirational," and that "people who don't tell stories about themselves lead vague lives." I found myself recalling Mr. Stienbaugh, my 10th grade biology teacher, and Mr. Apgar, my 7th grade math teacher. Why had these people come back to mind after all these years? During group discussions, we realized that kindness, openness, strength, and love are what count in the end. All the teachers we remembered had exhibited one or more of these qualities to some degree. I learned that the most important thing for a teacher is to try and be genuinely yourself

and appreciate your students in ways that express who you are.

Next, we both made our way the CALL SIG Annual General Meeting to meet up with some colleagues and hopefully make new friends, too. This was Jerry's first time, and he felt the atmosphere was quite like that of a bunch of old friends getting back together. Coordinator Tim Gutierrez got the meeting underway and Glenn Stockwell reminded everyone that the new SIG publication, *The JALT CALL Journal*, has now published two issues and invited those present to feel free to submit manuscripts. The CALL SIG finishes up this year with yet another successful conference, and planning has begun for next year's event at Sapporo Gakuin University in June of 2006. When they put out the word for proofreaders, we both volunteered. This will definitely be a good learning experience. Finally, it was time for dinner so we went out to the International Food Fair to enjoy some delicious food and a bit of tea, courtesy of M&P's Tea. Very refreshing!

Sunday morning started bright and early with *The Language Teacher's* annual staff meeting which was held on the 11th floor overlooking the beautiful Shizuoka Bay. After a bit of champagne, everyone got down to business under the leadership of Co-Editor Jacqui Norris-Holt. (Kim Bradford-Watts, the other Co-Editor, was busy pouring the champagne—an annual custom, I'm told). Jerry has been a proofreader for *The Language Teacher* for several years, while David has just started, yet we both felt like part of the group. It was good to see everyone's faces, and we were impressed with how seriously everyone takes their work.

Later Sunday morning, Jerry went to see David Nunan's plenary "Learning Styles & Strategies in the Classroom." He reports: "I went to this event in order to see a 'high-

level master' at work." Nunan began with the point that "learners need to be responsible for their education," something I agree with. He explained that there are four basic learning styles—concrete learners, analytical learners, communicative learners, and authority-oriented learners—and he pointed out what various studies had to say about this. He mentioned, "Good learners have a range of strategies that reflect their autonomy," and view language as a tool, not a subject. I found this distinction in line with my thinking and experience. He reminded me of the importance of knowing your learners, their learning styles, and how your teaching activities bias towards one of these. I will need to "audit" my teaching by looking at the activities I use, making sure that my preferred teaching styles take into account the various types of learning styles.

As a lover of all things Apple, I was intrigued to see "What can I do with an iPod in the classroom?" by Robert Chantrand. We use Macs at our school for video production, and I basically view them as "creativity liberation machines." I was hoping the presenter could give me some new ways of using Apple technology in my classes. Most of the presentation covered familiar territory such as what iPods are, how they are being used in education, their benefits, and difficulties. The presenter also went through some activities that made use of the iPod's ability to present sound and pictures, both music and spoken word using podcasts, recorded books, and photos. Now, with the recent release of the video iPod, there is also the possibility of using video content in the class. Overall, it seemed that you could conduct an English class just fine without an iPod, but if a teacher is into technology, those little machines offered up some exciting possibilities.

"Teaching and learning English meta-

phors” by Ramesh Krishnamurthy was an academic presentation about research on metaphors. The presenter, with a background in dictionary writing, made a lot of precise points on the nature of metaphors and how they are used in ELT textbooks. It turns out those metaphors are everywhere, but there isn't a lot on using or dealing with them. He presented several actual texts and showed how replete they were with obscure references. One conclusion was that there isn't much help for the teachers in decoding them. I'm currently using a book that contains lots of idiomatic and metaphorical expressions (*Innovations*, by Hugh Dellar), so this presentation provided some useful background information.

“Needs Assessment as Ongoing Teacher-Student Dialogue” by Kathleen Graves was a featured speaker workshop where we looked at needs assessment, something my school is beginning to get serious about. Again, I went in looking for inspiration and some practical

information I could take back with me. The workshop was very hands on. Teachers had to think about and create questions to assess student needs based on various criteria. This proved more difficult than at first glance, especially in the creation of “rubrics,” a common tool used in assessment these days. For example, how do you assess the learner’s level of inter-cultural competence? When we explored these topics, we could see clearly the complexity involved. In the end, I appreciated the opportunity to interact and ponder difficult questions with other teachers. The presenter gave several handouts with samples of what other teachers have been doing in this area. These proved highly useful, and will make good reading for my colleagues.

The JALT National Conference is such a grand event, it truly makes one wish they could be in two or more places at once. Next year's Conference in Kitakyushu will be something everyone in JALT must plan to attend. We hope to see you there!

From the Chalkface

Using Consciousness-Raising to Teach the Articles

Nic Farrow
Bunkyo Gakuin
Soka Women's College
Tamagawa Gakuen University

Summary

The misuse or omission of articles is a common problem for English learners. In this paper I present a lesson plan based upon consciousness-raising (C-R); the goal is for students to gain a general grasp of how articles work in English. Students should exhibit a substantial improvement of their understanding and usage of articles in just one lesson. Some follow-up techniques for consolidation are also given.

Materials & Preparation

Teachers using the examples provided need only prepare a few handouts in advance. However, some may want to make adjustments or create their own examples. The lesson handouts are available in *PowerPoint* format in the materials section of the CUE website: <http://allagash.miyazaki-mu.ac.jp/CUE>.

Time Required

One 90-minute period and time for optional follow up activities.

Rationale & Scope

C-R may offer an effective alternative to traditional methods of teaching article use.

According to Ellis (1994, p. 645) "Consciousness-raising provides a logical way of avoiding many of the pedagogical problems that arise from the teachability hypothesis." This lesson was designed with the goal of communicating a practical scheme in logical steps for students to gain a general grasp how articles are selected.

This lesson focuses primarily on the articles *a* and *the*, simply highlighting principles which govern the major problem areas which students are likely to encounter. The question of *zero articles* is not addressed, as my students' writing has consistently shown omission of articles to be a far greater problem than overuse.

This lesson introduces the use of *the* rather than *a* for the following:

- universally unique items (the sun)
- items unique to a situation (the floor, the grass, the kitchen)
- deictically obvious items (the guy over there)
- items of anaphoric reference (the guy we just referred to)
- items defined by a relative clause (the guy I met yesterday)
- superlatives (the nicest guy I know)

The article *a*, then, remains as the choice for singular nouns which have not been de-

fined according to the points above, reflecting its indefinite nature. Of course, various cases will occur in which the choice may be questionable, or governed by other, 'minor' rules. Nonetheless, this approach offers a comprehensible general guideline for use.

The Lesson Plan

Noticing/Discovery

a) Focus area: singular entities. This is a C-R 'noticing' exercise, using enhanced input of the target language to help students notice the 'rule' that universally unique entities usually take the article 'the.' This is easy to comprehend and sets the direction for the following exercises.

Provide the following examples on the board or with PowerPoint:

- The world is round.
- The sun is hot.
- Look at the sky.
- In the future we shall all have robots in our homes.

Elicit reasons why the sample sentences all use the article 'the.' The teacher can help by asking: *If you used these expressions, who would understand what you are talking about?* (Anyone, because everyone knows what you are talking about.)

b) Focus area: Deictic reference and superlatives. This is a C-R 'noticing' exercise, using enhanced input of the target language to help students notice the 'rule' that if both/all interlocutors in a certain situation can understand what is being mentioned, then 'the' is the article of choice.

Provide the following examples on the board or with PowerPoint:

Verbal interactions:

- A: "Be careful! The floor is wet."
- B: "Thanks."
- A: "Please close the door."

B: "OK."

- A: "I'll meet you at the station."

B: "OK. See you there."

- A: "Where's Wendy?"

B: "She's in the kitchen."

- A: "Pass me the kettle, please."

B: "Sure!"

- A: "What time is the last bus?"

B: "Ten forty-five."

- A: "The guy standing by the door is watching us!"

This could be done as a class or in groups, with students presenting their ideas after discussion. The teacher can help by asking: *In these cases, who knows precisely what or who is being discussed?* (BOTH, and the key is in the shared information in the situation—deictic reference)

Possible samples of written discourse:

- John was in his garden, sitting on the lawn.
- Stan parked his car in front of his house, and crossed the road.
- There are gift shops and restaurants on the first floor. (On a hotel's web site.)

This could be done as a class or in groups, with students presenting their ideas after discussion. The teacher can help by asking: *Why is 'the' definite in these cases?* (Because the reader can imagine the one item referred to, since there would likely only be one in each situation. In other words, the writer is indicating that the reader should know what is being talked about; he is not introducing information new to the reader.)

Provide the following examples on the board or with PowerPoint:

- Mt. Fuji is the highest mountain in Japan.
- A: "I need the cheapest film you sell."
- B: "Here you are, Sir."

- A: "The guy who was standing by the door is following us."

The teacher can help by asking: *What makes these definite?* (The superlative defines each one for both interlocutors, and the defining relative clause limits the man to one for both interlocutors.)

c) Focus area: anaphoric reference. This is a C-R 'noticing' exercise, using enhanced input of the target language to help students notice the 'rule' that first references to singular noun groups are usually preceded by 'a,' but subsequent references employ 'the.' Provide students with the following text:

Detectives Black and White had a big problem. There was a gangster... with a gun pointed at them! They were in a bar in downtown Chicago. The bar was almost empty, although the street outside was full of people. Black was reading a newspaper and White had a beer. The gangster was standing at the front door, and he did not look friendly. 'Am I the guy you were looking for, officers?' said the gangster. Joey Stun was the most dangerous gangster in the city. White slowly put the glass on the table, watching the man with the gun. Meanwhile, behind the newspaper, Black quietly switched on his radio to get help. 'Uh, Hi, Joey! It's a nice day, isn't it?' said White, nervously. 'Shaddup! I don't care about the weather! We are going for a nice little trip to the ocean! I'm going to give you a swimming lesson, with concrete boots!' said the gangster, 'Hurry up! Get up and go through the kitchen!' 'Can we get a drink before we go?' said Black. 'No way!' shouted Joey. Suddenly, they heard a siren. The noise was getting louder. The detectives looked out of the nearest

window. There was a patrol car about one block away. Joey pushed the detectives out the back door of the bar and into his car. White sat behind the steering wheel. He hit the gas, and the car was gone.

This could be done as a class or in groups, with students presenting their ideas after discussion. Using the passage provided, have students justify the use of 'a' and 'the' with each noun or noun phrase. Elicit how some of the 'the' noun phrases refer to previously occurring 'a' noun phrases.

Explicit Review & Discussion

The teacher needs to clarify these questions with the class, by elicitation.

'The' is for...?

'A' is for....?

If students are not able to explain, the teacher will have to do so.

The point to be established when we are writing or talking is *what do we want to tell our listener/reader when we choose 'a' or 'the'?* Do we want to say that he/she already knows about this—or are we introducing information we consider to be 'new'?

Given time, we might introduce for higher-level students some deeper insights. For example, if in our writing or speaking, we say *Mike went to the park and sat under the tree*, it indicates that it is a very small park! Why?

Find the Errors

Provide all students with the following handout. Students have to underline the errors and circle the correct uses of 'the' and 'a'. This could be done individually, in pairs, or in groups. This is a contrived, simplified text with limited pronoun use to minimize interference with the target language:

Kelly was a secretary in the small town in America. She was at home and she was having the bad day. She had a headache because she had been to the party the night before. She sat in a kitchen, and she drank the cup of coffee. She turned on the TV and began to watch a program about the man who was eating a big, fat hamburger. Kelly felt ill, so she found a remote control and changed to a new channel. On the new channel there was the aerobics program. Kelly picked up a phone and called her boss, but he was not in an office.

This can then be reviewed as a class, with students justifying their decisions.

From Definite to Indefinite

This exercise is for observation and reinforcement, though there can be elements of discovery regarding lexical and grammatical cohesion.

Provide all students with the following handouts. Students are to read a text and look for examples of this progression. Since students will also be exposed to the use of pronouns, this complication needs to be noted. Students should be encouraged to observe collocation of pronouns with their antecedents. Also, they should observe that similes (a lady/the woman) are subject to the rules of anaphoric reference.

A joke from the internet:

A rather simple young lady was in a casino in Las Vegas for the first time. She walked up to a soda machine, put in a coin and hit a button. Of course, a *Coke* came out of the machine. The lady looked amazed and looked in her purse for more coins. She started putting lots of coins in, and of course

the machine continued popping out soda cans.

Another customer walked up behind the woman. He watched her for a few minutes before stopping her and asking if he could use the machine.

The young lady shouted in his face: "GO AWAY! Can't you see I'm winning?"

A true story from the internet:

There is a real story from a small American town, like the hit movie "Home Alone."

In this story, while 13 year-old Ryan Hendrickson was home alone, a guy tried to break into his house. Ryan was watching a TV program, when he heard a noise. It was not loud, but he knew something was happening.

"I ran to the closet and got a baseball bat," Ryan said. "I went into the dining room with the bat, and I saw a guy cutting one of the windows with a knife. He put his left hand in first, and I was waiting for his right hand to come in. When he came through, I took the bat and I hit him as hard as I could. He dropped the knife and ran away. Then I called 911."

It is recommended that a little review/reinforcement take place in a following lesson by augmenting the lesson with exercises such as the ones below. These need not take up more than 30 or 40 minutes, and should serve to consolidate the principles covered by the main lesson.

Review 1: Extended Reading

PAIRS/INDIVIDUALS: Students identify and justify instances of articles in various selected reading passages. ('An' can be included here without much comment.)

Teachers can easily find materials for this purpose in many textbooks or other materials

appropriate to their students' level.

Review 2: Dictogloss

In a Dictogloss (described in detail by Wajnryb, 1990), the teacher reads the passage TWICE ONLY at a fair speed, with some pauses (faster than a dictation, slower than natural). This allows students to get the gist and write down substantial parts of the text. Then students work in groups to recreate the story with good lexis, grammar, etc. It is not important that the result be identical. The purpose is to have students use their linguistic knowledge and common sense to reproduce the story in grammatically acceptable English. I have found this technique useful to enhance grammatical awareness in a number of areas. The following text provides an example:

A man with a shotgun and a supermarket shopping cart walked into a bank. He shouted at the customers to lie down on the floor, and then he approached a window. The robber told the clerk at the window to give him all the money in her drawer. She handed him the money, but then she pushed him. He fell back into the cart, and the cart rolled across the bank and out the front door! A police car arrived and the officer arrested the man.

This passage has been reproduced with minor alterations with the permission of the author.

Reflections

Although I do not have data on the effectiveness of the lesson, anecdotally, I have to say it was successful. Even my least motivated students have shown interest in this subject. They seem to recognize this as a particularly puzzling problem in their communicative abilities. Many asked for

their post-lesson test results, even though I had not planned to present them. Preliminary results comparing written work before and after the lesson, and pre and post-lesson testing have been very promising. I have only used this on small classes, under 20, so I lack sufficient statistical evidence to prove its effectiveness. However, analysis of 18 students' diaries before and after the lesson, showed a 66% increase in the correct use of *the* though only a marginal improvement for *a* in their writing. Additionally, a remarkable side benefit was an apparent improvement in the students' use of determiners in general. Other determiner omissions (those I could not clarify from context) dropped by 66% in the post-lesson diaries, a figure that seems well worth further investigation. If anyone is interested in researching the effectiveness of this approach, I would be happy to hear from them. I have simple pre and post-tests prepared, as well as a grid for deeper writing analysis. To test the lesson's effect on spoken language may prove problematic, since it seems it would require a highly controlled experiment, although it may be possible in the form of a case study.

References

- Ellis, R. (1994). *The study of second language acquisition*. Oxford: Oxford University Press.
- Wajnryb, R. (1990). *Grammar dictation*. Oxford: Oxford University Press.

Conferencing

Steven Newman
Kinki College

Summary

Conferencing as described here is a warm-up/review activity using small groups as well as the whole class in a game format. Used at the beginning of a class, conferencing questions can be designed for students to review previous material, quiz themselves on trivia, preview new material, or explore background language knowledge.

Focus

Small group, cooperative review of previously covered/known material such as vocabulary, concepts, or trivia.

Level

Elementary and intermediate level college classes.

Purpose

In small groups, students practice speaking in a friendly yet competitive language game. Students interact and engage in cooperative groups. The format also activates students' knowledge of course material.

Procedure

1. Decide the focus of the conferencing warm-up, (e.g., grammar review, trivia, vocabulary, preview).
2. Prepare about five questions. The number of questions is flexible, but al-

low 2 to 3 minutes for each question. Questions should be clear and relevant; they should come from assigned readings, homework, previous lessons or future lessons. Provide a balance of serious, humorous, factual and open-ended questions to maintain a game show-like environment. Award points for correct answers. Ideally, the winners should receive some kind of bonus.

3. Arrange students into groups of three or four and set the task. Questions will be asked in succession. Groups will be given one minute to come up with an answer. All students should be prepared to answer since any one person may be called on to answer for the group.
4. Present questions and elicit answers. Feel free to award bonus points for effort and originality.

Variations

Students can make the class questions, though you should be sure to set strict boundaries (i.e., questions must be related to specific lesson or topic).

Rationale

With minimal planning, teachers can use this game-like technique to activate previous knowledge, process recent information, or review old material.