

Copyright
by
Margaret Enid Havelaar
2012

**The Report Committee for Margaret Enid Havelaar
Certifies that this is the approved version of the following report:**

**Improving Second Language Oral Production:
Teaching Implications from Recent Research**

**APPROVED BY
SUPERVISING COMMITTEE:**

Supervisor:

Elaine K. Horwitz

Veronica G. Sardegna

**Improving Second Language Oral Production:
Teaching Implications from Recent Research**

by

Margaret Enid Havelaar, B.A.

Report

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Master of Arts

The University of Texas at Austin

May 2012

Abstract

Improving Second Language Oral Production: Teaching Implications from Recent Research

Margaret Enid Havelaar, M.A.

The University of Texas at Austin, 2012

Supervisor: Elaine K. Horwitz

This work explores various methods teachers can use to promote high quality second language oral production. It consists of a review of empirical research and pedagogical implications related to the following factors: 1) Pre-task planning, 2) within-task planning, 3) task repetition, 4) task design, 5) formulaic sequences, 6) learner strategies, 7) form instruction, and 8) error correction. The work concludes with a consideration of issues within the literature and a brief summary of pedagogical implications.

Table of Contents

Introduction.....	1
Background	1
Goals for Oral Production	1
Rationale for Topics Addressed in This Report.....	2
Useful Factors in Oral Production	4
Introduction.....	4
Pre-Task Planning	4
Conclusion	7
Summary of Pedagogical Implications	8
Within-Task Planning	8
Conclusion	11
Summary of Pedagogical Implications	12
Task Repetition	12
Conclusion	14
Summary of Pedagogical Implications	15
Task Design	15
Conclusion	18
Summary of Pedagogical Implications	20
Formulaic Sequences	20
Conclusion	23
Summary of Pedagogical Implications	24
Learner Strategies	24
Conclusion	27
Summary of Pedagogical Implications	28
Form Instruction.....	28
Conclusion	31
Summary of Pedagogical Implications	32
Corrective Feedback	32

Conclusion	38
Summary of Pedagogical Implications	39
Discussion	40
Literature Issues	40
Conclusion	42
References	45

Introduction

BACKGROUND

I became interested in the topic of student oral production in second language classes after researching French immersion programs in Canada. As a former French immersion student, I have a high opinion of the program and of the concept of language immersion programs in general. French immersion programs in Canada have produced graduates with strong listening and reading skills and the ability to confidently express thoughts and opinions, formulate arguments, and analyze issues in their second language. However, a number of studies have suggested that French immersion graduates produce a version of French suggestive of a fossilized interlanguage. This possible oral proficiency gap motivated me to investigate oral production research to search for clues that could help my teaching.

Although the topic of this Report arose from my interest in French immersion programs, the literature I have reviewed encompasses diverse educational contexts and its pedagogical implications can be applied well beyond the context of Canadian French immersion programs.

GOALS FOR ORAL PRODUCTION

The goals for oral production in a second language are many. As a teacher, my primary goal in assisting my students with their oral production relates to their ability to convey meaning through speech. This goal is largely due to my interest in teaching methodologies such as communicative language teaching and task-based language

teaching (TBLT), both of which encourage meaningful communication, as well as my personal dedication to developing proficiency in second language classrooms. Several factors explored in this Report contribute to students' ability to effectively convey meaning. Firstly, better grammatical and syntactic accuracy lead to fewer errors, and therefore more comprehensible output. Secondly, grammatical, syntactic, and lexical complexity allow students to use more advanced language for a variety of interactions. Thirdly, increased fluency allows students to produce language in real time without non target-like pauses and hesitations. Finally, pragmatic concerns such as not offending interlocutors can assist in the conveyance of meaning. Pronunciation remains a concern for many students, however I regard it as a separate issue for the purposes of this Report and have instead focused on other measures of oral production.

The overarching goal for the development of my students' oral language skills is to promote long-term second language acquisition whereby students are able to transfer skills gained in the classroom to new settings and contexts. Without long-term benefits to students' ability to communicate verbally outside the classroom, improvements in accuracy, complexity, and fluency will only help them achieve better grades in assessment situations or help them communicate immediately in the classroom.

RATIONALE FOR TOPICS ADDRESSED IN THIS REPORT

In this Report, I have chosen to discuss eight factors with the potential to influence student L2 oral production. I had originally hoped that a textbook author would provide me with a list of factors on which I could build my research plan, but although I found inspiring advice in many second language teaching methodology textbooks, I could not isolate a clear set of factors for my purpose. Eventually I stumbled upon literature

related to TBLT that helped guide me toward the eventual eight topics covered in this Report. Ellis' (2005) and Bygate, Skehan, and Swain's (2001) compilations of TBLT-related studies provided the most guidance on possible topics, and I have reviewed several of the studies from these compilations in this Report. Building upon this foundation, I continued my search for empirical studies, eventually settling on eight topics that were well enough represented in the literature to warrant a review. Of the final eight topics, four are directly related to TBLT (pre-task planning, within-task planning, task repetition, and task design) and four other factors warranted an exploration due to their obvious pedagogical implications (formulaic sequences, learner strategies, form instruction, and error correction).

Some factors such as topic familiarity or individual differences were interesting, but there simply were not enough empirical studies directly related to oral production on these topics to allow for a thorough review. Some of the topics, notably error correction and form instruction, offered too many studies for me to review all of them. In these cases, I attempted to select the most relevant or notable studies from the past decade. I decided that the many studies demonstrating computer assisted language learning's (CALL) potential for assisting oral production were outside the scope of this Report because they often involved educational settings where there was no teacher. My exclusion of CALL does not imply that I think it has no potential for promoting more target-like oral production. On the contrary, it appears to be a promising area of research.

I generally reviewed studies from the past decade, but occasionally delved into a study from the late 1990s. I also limited my review to studies that explicitly measured oral output. I excluded the wealth of research on second language acquisition (SLA) theory from this Report, choosing to focus on pedagogy rather than L2 acquisition processes for the most part.

Useful Factors in Oral Production

INTRODUCTION

In this Report, I will explore eight factors that could help promote more accurate, complex, and fluent student oral production, with the goal of increasing students' ability to convey meaning through speech. These factors are 1) pre-task planning, 2) within-task planning, 3) task repetition, 4) task design, 5) formulaic sequences, 6) learner strategies, 7) form instruction, and 8) error correction. In each of these sections I will review empirical research on the topic and provide a list of pedagogical implications based on the results of the studies and researchers' conclusions. Teachers can incorporate these suggestions into their own teaching as appropriate for their educational setting in order to promote high quality oral communication in their students.

PRE-TASK PLANNING

Pre-task planning, also often referred to as strategic planning, has garnered a great deal of attention in the last two decades. This may be due to the interest in researching Task-Based Language Teaching (TBLT). Pre-task planning can take many forms and includes any activity that prepares students for a task. Rather than memorizing or even planning detailed formulations, "learners plan propositional content and isolated chunks of language to encode it" (Yuan & Ellis, 2003, p. 7).

Much of the research for pre-task planning allows for ten minutes of planning time before students are asked to produce an oral task (Kawauchi, 2005). Kawauchi (2005) allowed ten minutes of planning time in her study comparing native Japanese-speaking students of English at three different proficiency levels: low EFL, high EFL,

and advanced ESL. While all groups benefitted from pre-task planning, the high EFL group benefited the most, attaining the fluency level of the advanced ESL group. The high EFL group also gained the most in terms of complexity and accuracy. For comparison, when this same group performed an unplanned task, they were significantly less fluent. Curiously, the advanced ESL group performed less fluently on the planned task than on the unplanned task, using more repetitions in the former. Kawauchi notes that this result contrasts with much of the research on similar student populations.

Yuan and Ellis (2003) also allowed for ten minutes of planning time in their study of English majors at a Chinese university's international business school. Contrary to most other similar research, they found that pre-task planning had little effect on fluency compared to a control group. They also found that it had little effect on lexical or syntactic variety and only a slight effect on accuracy, but there was a positive effect on syntactic complexity. The fluency result may be related to the fact that the authors controlled for within-task planning by limiting the amount of time for students to perform the task, which equalized any positive effects the pre-task planning may have provided (Yuan & Ellis, 2003). When compared to a group that had time for within-task planning, the pre-task group outperformed them on fluency measures but not on accuracy measures.

Tavakoli and Skehan (2005) provided five minutes of pre-task planning to participants in their study. Iranian elementary and intermediate EFL students all improved on measures of accuracy, complexity, and particularly fluency. Elder and Iwashita (2005) examined the effects of only three minutes of planning on oral tasks in an assessment situation. They found little difference between the treatment group and the control group in terms of accuracy, fluency, or complexity. Unfortunately, their control group was given 75 seconds of planning time, leaving only an extra 105 seconds to the

treatment group. Furthermore, students were given plenty of time to perform the task, which means that they had time for within-task planning during the task. Overall, studies that provide students with more time for pre-task planning have produced more encouraging results.

In addition to the amount of time provided for pre-task planning, studies that compared different types of planning methods suggest that pre-task structure matters. For example, Kawauchi (2005) looked at the effect of two types of planning (reading versus writing) on oral narrative tasks. In terms of the type of planning, both writing- and reading-focused pre-planning were effective, with a slight advantage for reading likely due to the difficult lexical items that were recycled during the oral narratives. Comparing teacher-led planning to solitary and student group-led planning, Foster and Skehan (1999) looked at college students from a variety of language backgrounds learning English in the United Kingdom. They found significant effects on accuracy when teachers led the pre-task planning and advantages for measures of fluency and complexity as well. Solitary planning was also successful, benefitting complexity, fluency, and turn length. Finally, the results for group planning were similar to the control group. The teacher-fronted pre-task planning appears to provide the most benefits to students in terms of oral performance, with solitary planning a close second. Foster and Skehan (1999) also compared planning for language to planning for content and found little difference in output.

Finally, rehearsal can be considered as a specific type of pre-task planning (Ellis, 2009). In his review of rehearsal studies, Ellis (2009) found evidence in favor of rehearsal, particularly as it relates to fluency and complexity, although there is likely “no transference of the rehearsal effect to a different task, even when this is the same type as the original task” (p. 476). Ellis recommends that teachers provide feedback to students

in order to affect acquisition. Therefore, although rehearsal may be useful for students looking for higher scores in assessment situations (where rehearsal is possible), additional interventions may have more long-term benefits.

Most of the nineteen studies reviewed by Ellis (2009) found a positive effect for pre-task planning on fluency, especially in less complex tasks and when ten minutes of planning time was built into the study. The research reviewed in this Report contains a few surprises but at least moderately corroborates Ellis' findings for fluency. On complexity measures the results are much more mixed, although improved grammatical complexity appears to be a common theme in the literature (Ellis, 2009). Ellis also found mixed results for accuracy measures. They are much more positive in the studies I have reviewed in this Report, with the exception of Kawauchi (2005). Ellis (2003) suggests that there may be a trade-off effect, suggesting that "when learners plan they have to choose what aspect of production to focus on; focusing on fluency and complexity is at the expense of accuracy and vice versa" (p. 133).

Conclusion

It should be noted that most of the studies of pre-task planning test for immediate results rather than long-term results. Although delayed post-tests are difficult to implement and even more difficult to control for other influencing factors, more long-term results would offer stronger evidence of skill development beyond immediate classroom or assessment situations. Elder and Iwashita (2005) express their concern about claims "to test language communicatively or, in other words, to mirror the demands of real world language encounters in assessment contexts" (p. 235).

Given the post-test caveat noted above, it may be difficult to offer any clear advice to teachers based on the studies I have reviewed. However, pre-task planning does

appear to be promising for increasing fluency at least, if not also complexity and accuracy. As with all of the topics covered in this Report, context and task design are major factors in the effectiveness of any task condition. Elder and Iwashita's (2005) study suggests that students either need training on how to use their pre-task planning time wisely or more planning time (more than three minutes) for it to have a positive effect on oral tasks. The results from Foster and Skehan's (1999) study suggest that teacher-led planning is most effective for output during tasks, but there may be untested trade-offs in their study, such as a decrease in learner autonomy.

Summary of Pedagogical Implications

- Teachers should allow five to ten minutes of pre-task planning time before a task.
- Pre-task planning involving reading or writing are both effective.
- Teacher-led and solo student planning may be more effective than student group planning.
- Teacher should provide feedback to students after a rehearsal in order to promote long-term acquisition.

WITHIN-TASK PLANNING

Within-task planning, also referred to as online planning, can be defined as the planning that takes place while students are engaged in a speaking task. Although some definitions include students' "pre-production and post-production monitoring of their speech acts" (Ellis & Yuan, 2005, p. 6), researchers often operationalize within-task planning by controlling for pre-task planning time, and simply provide plenty of time for students to perform the task. According to Ellis (2003), "when learners use the time at their disposal to attend to formulation and to monitor the use of their grammatical

resources their production becomes more accurate” (p. 128). Students may also use the time to attend to other output elements, including complexity, content, and meaning. Fluency, defined by Ellis (2009) as “the capacity to use language in real time, [and] to emphasize meanings” (p. 475), however, is very unlikely to increase as a result of within-task planning.

In their study of English majors at a Chinese university, Yuan and Ellis (2003) observed the effects of strategic and within-task planning on various output measures. When provided time for within-task planning, students’ grammatical accuracy on a narrative task improved significantly, and their grammatical complexity increased compared to the control group, likely because students had time to access their explicit grammar knowledge. There were no benefits, however, to fluency or lexical variety. In a study featuring similar participants to Yuan and Ellis (2003), Ellis and Yuan (2005) also found that the opportunity for careful within-task planning resulted in increased complexity and accuracy, though it had no effect on lexical variety. The authors explain that “in the case of speaking (whether rapid or careful) there is little time for conceptualisation as learners must necessarily engage with the process of within-task production” (p. 191). They also found that the opportunity for within-task planning resulted in slight disfluencies.

There have been several studies pointing to the disfluency that “naturally results from learners’ engagement in careful online planning” (Ahmadian & Tavakoli, 2011, p. 38). This disfluency may occur when students access their rule-based L2 system during within-task planning (Ahmadian & Tavakoli, 2011). To attempt to remedy the disfluency associated with within-task planning, Ahmadian and Tavakoli (2011) studied the effects of combining two task factors, careful within-task planning and task repetition, on post-secondary intermediate English learners in Iran. As with Yuan and Ellis (2003) and Ellis

and Yuan (2005), Ahmadian and Tavakoli (2011) found that careful within-task planning led to significantly better accuracy and complexity. Adding the task repetition condition to within-task planning made little extra difference to accuracy, though it should be noted that task repetition without time for within-task planning also produced much greater accuracy. The combined task repetition and within-task planning condition also led to an “exponential increase” in complexity and greater fluency than the control group (Ahmadian & Tavakoli, 2011, p. 55). In summary, adding task repetition to a controlled within-task planning task may compensate for the disfluency associated with within-task planning. The combination of these two factors led to gains in all three language production measures in this study: fluency, accuracy, and complexity.

Ellis (2005) reasons that accessing the rules of a language takes a toll on working memory (WM), leading to disfluency. Yuan and Ellis (2003) argue that within-task planning “allows time for the central executive of WM to operate and thus enables learners to search their LTM [long term memory] for grammatical encodings” (p. 7). Ben Maad (2010) also suggests that the removal of time pressure eases the burden on WM, allowing speakers to use a holistic processing mode rather than an analytic processing mode. In an effort to know more about the relationship of working memory to oral production under within-task planning conditions, Ahmadian (2012) tested the WM capacities of forty Iranian EFL students and asked them to perform oral tasks under within-task planning conditions. He found no correlation between the oral production of students with high WM capacities and better complexity scores, suggesting that despite individual differences related to WM, many students can benefit equally from within-task planning in terms of complexity. On the other hand, Ahmadian found statistically significant positive correlations between WM capacity and both accuracy and fluency. It is interesting to note that at least some students produced fluent language after a within-

task planning task. However, we have no comparison, as there was no control group. We only know that students with higher WM capacity produced more fluent and accurate language than those with less WM capacity.

Conclusion

Ellis (2009) claims that there are too few studies of within-task planning to come to any solid conclusions about its effectiveness. Given that the studies reviewed here (from before and after 2009, when Ellis conducted his review) have largely corroborated each other regardless of educational context, it is fair to conclude that teachers in at least some contexts can expect that the opportunity for within-task planning will help foster complexity and accuracy in their students' oral production, albeit perhaps at the expense of fluency.

Unfortunately, within-task planning may have little practical application. Real-world situations in the target language may offer few opportunities for careful within-task planning, and fluency may be a more important factor in non-classroom situations. Teachers should keep these caveats in mind, but should also consider that within-task planning may be a useful teaching tool. Because of the lack of longitudinal research on the benefits of within-task planning, insights into its usage in the classroom must remain speculative. Allowing time for careful within-task planning could be an effective confidence builder, but it could also serve little purpose in the long term. However, Yuan and Ellis (2003) do not believe that the classroom must replicate real life, and suggest that within-task planning could prepare students for eventual communication beyond the classroom.

Summary of Pedagogical Implications

- Teachers should provide students with plenty of time in which to perform an oral task in order to promote accuracy and complexity. Time for within-task planning may be combined with task repetition to promote fluency.

TASK REPETITION

Ahmadian and Tavakoli (2011) argue that task repetition can result in improved oral output due to the brain's control mechanism requiring less processing power during repetition. Attentional resources can thus be diverted to processing beyond the initial conceptualization to focus instead on formulation and articulation (Ahmadian & Tavakoli, 2011; Bygate, 2001; Lynch, 2007). At this point, "some traces of the forms that [students] have used in their first performance may be available to them" (Ahmadian & Tavakoli, 2011, p. 56).

Ahmadian and Tavakoli (2011) operationalized task repetition as a repeated oral task with a one-week interval. They found that task repetition had a positive effect on fluency and complexity. Because the main goal of their study was to observe the effects of task repetition combined with careful within-task planning, they discuss the fact that the two groups of students who engaged in careful within-task planning did approximately equally well in terms of accuracy, regardless of whether task repetition was added as a condition or not. Thus it would appear that the effect of task repetition on accuracy is limited, and that careful within-task planning has more of an effect on accuracy than does task repetition.

Bygate (2001) operationalized task repetition by allowing a full ten-week interval between tasks. In his study of English learners at a British university, he found positive results for fluency and complexity, but not for accuracy, after students repeated the same

task. Two different types of tasks, narrative and interview, yielded approximately the same results. For repetition of different tasks of the same task type, the gains were less obvious, but Bygate suggests that “there may be a residual gain which can be found from task-type exposure when a specific task is repeated” (p. 44).

In his study of sixteen graduate students from ten countries learning English, Lynch (2007) explored transcription as a form of task repetition. He believes that rendering students’ speech visual through transcription can help them notice problem areas, thus propelling them toward higher quality output, particularly with regard to form. Students who performed their own transcriptions achieved higher quality oral output one month later compared to students for whom the teacher performed the transcription.

Mennim (2003) also explored transcriptions with undergraduate students of English in Japan. His positive results for output corroborate those of Lynch (2007), although the follow-up oral task in this case was an exact replica of the original oral presentation with fewer mistakes. In fact, the initial performances in this study were rehearsals, and Mennim did not control for students simply memorizing their presentations after transcription and error correction.

While the higher quality of language emerging from task repetition could be motivating and anxiety relieving for students, repetition also has the potential to be boring, and perhaps even demotivating. Lynch and Maclean (2001) observe the effects of immediate task repetition, but find a creative way of designing task repetition to ensure that it remains interesting for students. In fact, the participants “did not perceive the task to be repetitious” (p. 159), quite a feat given the number of times they repeated the same task. The participants in this study were students of English, all of them oncologists and radiotherapists taking a class to help them prepare for presenting papers in English at conferences. The students presented posters at an information session, communicating

essentially the same information to a series of visitors. Thus, “the learners’ changes to their English came about not in response to external intervention before or during the task, but in pursuit of communication with a series of different visitors” (p. 158).

Not all students will notice their own improvement after task repetition. In Lynch and Maclean’s (2001) study, weaker students in particular were less likely to notice improvements. Lynch and Maclean (2001) conclude that “teachers should design and include post-task activities to help them to monitor ways in which their performances had in fact become more accurate, but which they were too preoccupied to notice in the heat of communication” (p. 157). Ellis (2009) suggests that feedback after the initial performance can help students improve their L2 performance.

Conclusion

Samuda and Bygate (2008) defend task repetition as a worthwhile teaching tool because “we frequently repeat discourses in everyday life, and this has been associated with language development in children [. . .] and at least anecdotally for adults” (p. 114). Furthermore, practicing the same task more than once may help with new tasks of the same type because they will contain enough similarities to reduce the cognitive load for speakers (Bygate, 2001).

The benefits of task repetition do not imply that spontaneous language should be ignored in the classroom. Spontaneous unrehearsed language remains an important part of classroom communication, “and yet to provide speaking practice only under these conditions runs the risk that learners will constantly be improvising, constantly experimenting with new forms, but also constantly doing so while having to pay some considerable attention to the content of what they want to say” (Bygate, 2001, p. 44). Bygate (2001) suggests that task repetition allows students the opportunity to practice

communication in a different manner, and that by providing time for both rehearsed and spontaneous language in the classroom, students can focus on many aspects of language.

Summary of Pedagogical Implications

- Task-repetition in the form of student presentations is beneficial when the need for students to repeat information to different people arises naturally.
- Feedback or other types of post-task activity after task repetition help students reflect on their improvement.
- For some contexts, transcription of oral presentations can be helpful for improving oral production.

TASK DESIGN

Teachers have the option of using many types of tasks in the classroom. These can include narratives, interviews, arguments, and instruction-giving, problem-solving, and decision-making tasks, to name a few. Researchers have found varying benefits to oral production depending on the task. For example, Aliakbari and Jamalvandi (2010) found role-play to be challenging and motivating for the 60 Iranian students of English in their study. These students were more fluent, accurate, and used more complex vocabulary than a control group on a post-treatment speaking test. Tasks can also be designed to require certain types of language, such as requests, refusals, or opinions. Different tasks can encourage specific learner behaviors such as more self-repair (Gilabert, 2007), or increased negotiation of meaning and interaction (Gilabert, Barón, & Llanes, 2009). Finally, tasks can be manipulated to render them more or less complex, more or less structured, and can also be sequenced to maximize the benefits to student

output. The literature reviewed here offers suggestions for teachers on how to approach task design, including assessments.

Several researchers have looked into the impact of task structure on oral performance. Wigglesworth (2001) focused on assessment design. She found that more structured tasks were easier for students, particularly for information-gathering tasks, and somewhat less so for negotiation tasks, but she does not explicitly say that easier tasks result in better oral production. Wigglesworth explains that scaffolding on which students can build language helps reduce cognitive load. The main pedagogical implication to be gleaned from this study is that assessment design should be carefully considered. Student performance will vary depending on task structure, among other factors that may be more obvious to teachers, such as complexity or difficulty.

Tavakoli and Skehan (2005) compared the benefits of structured and unstructured tasks for Iranian students at an English language school in Tehran. They found significantly less output accuracy with the two unstructured tasks in their study, and significantly more output complexity in one of the structured tasks. The authors could not explain the result for complexity given that the other three tasks in the study (one structured and two unstructured) all had very little effect on complexity.

Skehan and Foster (1999) and Tavakoli and Foster (2008) also looked into task structure, focusing specifically on clear inherent sequential structure in narrative tasks. The participants in Skehan and Foster's (1999) study were university students in the United Kingdom representing various L1 backgrounds. The authors found that clear linear sequential story structure positively influenced fluency, but that it had little impact on complexity or accuracy. Tavakoli and Foster's (2008) study shows competing results; clear linear sequential structure led to better accuracy but had little effect on fluency. As with Skehan and Foster (1999), the impact on complexity was minimal.

Tavakoli and Foster (2008) also looked into task complexity and found that it too had little effect on oral fluency, but yielded positive results for output complexity. Their study included a comparison of the effect of the same task on similarly proficient learners in London and Tehran. They found that the London students consistently achieved higher complexity and more diverse vocabulary than the Tehran students. Robinson (2001) also looked into task complexity, this time with Japanese university students of English. He found that more complex tasks led to more lexical variety and that simpler tasks led to more fluency.

In addition to task complexity, Robinson (2001) also explored task sequencing. His study reinforces the notion that teachers should sequence tasks from simplest to most complex in order to promote accuracy and fluency. He adds that “sequencing tasks on the basis of their cognitive complexity is to be preferred over sequencing decisions based on task difficulty or task conditions” (p. 51). Although the terms *complexity* and *difficulty* are often used interchangeably, Robinson views them differently. Complexity, he writes, is “the result of the attentional, memory, reasoning, and other information processing demands imposed by the structure of the task on the language learner” (p. 29), whereas he describes difficulty as being related to individual variables in learners, such as those related to ability and affective issues. Of course, task complexity and difficulty often go hand in hand (Robinson, 2001). And, even when sequenced in a different order, the ratings of difficulty for the most complex tasks remained the same in this study.

Bygate (1999) explored the effects of two different tasks, a narrative task and an argument task, on complexity measures. Based on the results from his study of Hungarian high school students learning English, he suggests that argument tasks tend to lead to the use of more verb groups, more individual verb forms, and more sequential language, and that none of these benefits come at the expense of plenty of improvised speech. The

narrative tasks led students toward longer turns and “linguistically denser talk” (p. 206). Bygate suggests that narrative tasks may be more difficult to sustain over a longer period of time than argument tasks, due to more complex processing.

In his study of Japanese undergraduate students of English, Taguchi (2007) looked beyond the usual complexity, fluency, and accuracy measures to explore appropriateness, an especially important consideration in Japanese L1 norms where polite and informal language are more dissimilar than they are in English. Two pragmatic speaking tasks “measured the learners’ ability to understand situational information and to perform two speech acts, requests and refusals, in role plays” (Taguchi, 2007, p. 119). The more difficult task featured a situation with a high power difference, social difference and degree of imposition, while the other featured the inverse characteristics. The low proficiency students had more trouble with the difficult task, resulting in lower appropriateness scores. The author suggests that pragmalinguistic factors matter when planning a task, particularly for the context of this study.

Conclusion

Based on the studies reviewed here, it is fair to suggest that task design usually impacts the type and quality of student language. Ellis (2003) summarizes the task conditions most likely to lead to specific qualities in oral production. He suggests that task design may not have as much impact on accuracy compared to fluency and complexity measures, although there is some suggestion that when tasks are combined with pre-task planning time, the effects for accuracy are greater. The factors that may elicit more accurate language are a lack of contextual support, tasks with no definite answer, and structured tasks (Ellis, 2003). Structure had little impact on accuracy in Skehan and Foster (1999), but did have an impact in Tavakoli and Foster (2008). As

Robinson's (2001) study demonstrates, proper sequencing of tasks can also positively affect accuracy.

There is much more compelling evidence for task variables being able to elicit more complex or fluent language. For complexity, Ellis (2003) suggests less contextual support, more elements in a task, and tasks that allow many solutions as opposed to only one. Tavakoli and Skehan's (2005) study suggests that for structured versus unstructured tasks, the effect on complexity remains unclear, and Skehan and Foster (1999) saw no effect of task structure on complexity. Tavakoli and Foster (2008) corroborate Ellis' (2003) summary for complexity measures, in that a more complex task (i.e., one with more elements) can lead to more complex language. For fluency measures, Ellis (2003) suggests that tasks with more contextual support and a clear structure, tasks that have a single correct solution, and tasks that "pose a single demand" (p. 127) may have the most effect. Tavakoli and Foster (2008) found little effect of task structure or task complexity on fluency. Finally, proper sequencing of tasks may be beneficial to fluency (Robinson, 2001).

Of course, the utility of such advice is limited given that there appear to be trade-off effects for many of the tasks. Different tasks "may predispose learners to engage in certain types of production but they cannot guarantee them" (Ellis, 2003, p. 127). Tavakoli and Foster (2008) also suggest that teachers can manipulate tasks to promote specific goals in oral production. As the quality and volume of research increases, the outcomes become more predictable, they claim. Bygate (1999), on the other hand, maintains that teachers probably cannot "plan and implement tasks as if these could determine learners' language use" (p. 207). Yet he concedes that teachers can influence language use by selecting different task types, and can also influence output by rendering tasks less or more complex.

The summary of beneficial task types for each output measure should provide some guidance for teachers who have specific goals in mind for their students. It also suggests that to promote all three goals of fluency, accuracy, and complexity, a variety of tasks will be necessary. Unfortunately, none of the studies reviewed here measured long-term acquisition of a second language. Skehan and Foster (1999) feel that it is nonetheless reasonable to extrapolate the results from immediate post-tests to longer-term development, although they offer little in the way of justification for this view. Until research addresses this gap, the interpretation of immediate results must be sufficient for influencing task design.

Summary of Pedagogical Implications

- More structured tasks and less complex tasks may be easier for students and promote more target-like output.
- Teachers should sequence tasks from least to most complex.
- Teachers should implement a variety of task types to maximize the number of aspects of language to be practiced.
- Teachers should consider pragmalinguistic factors when planning certain tasks.

FORMULAIC SEQUENCES

Formulaic sequences can include collocations, idioms, clichés, and phrases used in social routines. In theory, these chunks of speech become memory-based language over time and therefore more easily accessible during a speech act, eventually requiring less mental processing than spontaneous language (Stengers, Boers, Housen, & Eyckmans, 2011; Taguchi, 2008). This description suggests a strong connection to automaticity. Ben Maad (2010) argues that the use of formulaic sequences can function

“as a time-buying strategy” to give students “extra time to monitor their output” (p. 600). Additionally, interlocutors may perceive speakers who use more formulaic sequences to be more native-like and fluent (Stengers et al., 2011). In short, there are several potential benefits to students incorporating formulaic sequences into their speech.

Ben Maad (2010) studied the effects of different task conditions on fluency, accuracy, and complexity. His results, based on the observation of Tunisian undergraduate students learning English, suggest that students use different processing modes for different task conditions. Students’ use of formulaic language was associated with the holistic processing mode whereas students in time pressure situations made use of analytic processing at the expense of formulaic language and its corresponding fluency and accuracy. Instead, the time pressure task resulted in greater complexity. These results suggest a trade-off effect as a result of two different processing modes. Ben Maad advises teachers to spend time on formulaic sequences due to the advantages of holistic processing associated with them. Furthermore, quick access to pre-fabricated formulaic sequences may help ease anxiety and motivate students as their fluency increases (Ben Maad, 2010).

As with all of the topics addressed in this Report, the context of any study should be considered when attempting to extrapolate results to other contexts. The importance of formulaic sequences to fluency may depend on the specific target language. In their study of Belgian Dutch-speaking university students similarly proficient in their respective target languages of Spanish and English, Stengers et al. (2011) found that the number of formulaic sequences was more positively correlated with oral proficiency ratings for the English learners than for the Spanish learners. The authors speculate that more analytic languages, including English, lend themselves to mastery of formulaic sequences more

easily than synthetic languages like Spanish. In addition, the authors consider English to be much more idiomatic, and therefore more formulaic, than Spanish.

In another study of English majors at a Belgian university, Boers, Eyckmans, Kappel, Stengers, and Demecheleer (2006) found that when students were taught to notice sequences in authentic listening and reading materials, they not only used more formulaic sequences when speaking than students who were taught with a more traditional focus on grammar and vocabulary as separate aspects of language, but also were perceived to be more proficient in terms of fluency and range of expression. Although these results are promising, the long-term effects of teaching noticing are unclear. The authors point out that the students in this case may not have actually built up their repertoire of sequences. Rather, they were taught to notice them, and recycled sequences they had only just read during the oral retellings that formed the basis of their fluency ratings. This certainly does not invalidate the utility of teaching students to notice structures, but it does suggest that additional interventions may be necessary for longer term gains.

Rather than teaching students to notice sequences, Taguchi (2008) exposed students of Japanese at a U.S. university to forty specific grammatical chunks through communicative drills and memorization of dialogues containing the chunks. She found that the students significantly increased the complexity of their oral production. They not only incorporated the new chunks into their oral production, but also “learned grammatical chunks to construct complex and extended syntactic turns out of simple, small chunks” (p. 146). In short, explicit knowledge of specific sequences allowed students to use them as building blocks on which they could create more complex language. Curiously, there were no gains for fluency in this study, as operationalized by speech rate and pause length.

Conclusion

Empirical research at this point provides no clear answers on how to teach formulaic sequences, or which ones to teach. Gatbonton and Segalowitz (2005) have started the conversation by encouraging teachers to focus on “utterances that have clear pragmatic functions” (p. 333). Their ACCESS (Automatization in Communicative Contexts of Essential Speech Segments) methodology, based in communicative language teaching, includes three phases. Integrated within the first phase is a pre-task and a main task in which “learners engage in a task or tasks in which functionally useful utterances are used and elicited naturally and repeatedly” (p. 329). Repetition of the formulaic sequences is essential to them, though it should be noted that they are not referring to the type of repetition in Taguchi’s (2008) study. Rather, repetition should be essential for students to reach the goal of a task. In comparison to the methodology in Taguchi’s (2008) study, Gatbonton and Segalowitz offer a much more appealing method by which to teach formulaic sequences. The drills and memorization outlined in Taguchi’s study have the potential to be boring and perhaps even demotivating for students. Taguchi recommends that the practice outlined in her study “be supplemented by frequent free conversation practices in which learners practice using a group of learned chunks in tandem to construct discourse” (p.151). Teachers interested in more communicative or task-based instruction may find ways to incorporate this advice into their pedagogy.

Currently lacking in the literature on formulaic sequences are details on how chunks of language are stored in the brain and how they are accessed. This gap renders it difficult to provide more solid advice to teachers without more research into both second language acquisition and teaching methodologies. Despite these issues, there are some pedagogical implications to be gleaned from the research examined here. It suggests that formulaic sequences are an important aspect of oral proficiency and should not be

ignored in the classroom. They may be particularly useful for advanced learners who have reached a plateau, “since acquiring a rich and varied formulaic language would enable [them] to reach a native-like level” (Ben Maad, 2010, p. 600).

Summary of Pedagogical Implications

- Teaching students to notice formulaic sequences may prove beneficial.
- Teachers should integrate pragmatically useful sequences into tasks where students will use them naturally and repeatedly in order to meet a goal.

LEARNER STRATEGIES

Teachers can help students learn a language by steering them toward the development of effective strategies. Learners either consciously or subconsciously “use methods or approaches appropriate to their needs to memorize and understand the target language, to monitor and evaluate learning progress, to compensate for lack of knowledge and to interact with other people” (Chou, 2011, p. 274). Encouraging students to use such strategies is much less controversial than deciding how teachers can foster effective strategy use. In this section I will explore empirical research on strategies and its pedagogical implications for assisting students with their oral production. In recent years, several researchers have attempted to test the effects of explicit strategy instruction, also referred to as metacognitive strategy instruction, on students’ oral performance. Lam (2009) explains that metacognitive strategies “oversee the general learning process by enabling the learner to think ahead of the goal and demand of the learning task, to plan for some action to tackle the task, and to assess how well one has done the task” (p. 130). Although researchers have not always focused on the same strategies, overall they have designed instruction to help improve students’ oral

performance on a variety of measures, including confidence, pronunciation, and fluency, among others.

Lam (2009) observed two classes of English students in Hong Kong that were explicitly taught seven metacognitive strategies over the course of five months. The students who received the instruction eventually outperformed the control group in a group discussion task wherein raters judged students on their “confidence when handling the task, cooperation/mutual help in conducting the task, and general effectiveness in fulfilling the requirements of the task” (p. 137) as well as their pronunciation, vocabulary, and grammar use. In the class discussion, the students clearly did not make use of all of the strategies taught, even though they self-reported having used all seven of them. Lam explains that students may have declarative knowledge about strategies after training, but that this awareness has not yet been proceduralized.

In two similar studies also examining explicit metacognitive strategy instruction, Nakatani (2005) and Nakatani (2010) observed Japanese college students in an EFL setting. Nakatani’s 2005 study focusing on oral communication strategies found that students improved significantly compared to a control group on seven factors related to fluency, ability to interact with interlocutors, and flexibility. Through explicit teaching, a reference list of strategies, and the use of strategy diaries, students “came to make longer utterances and use more achievement strategies, such as modified interaction, modified output, time-gaining, and maintenance strategies” (Nakatani, 2005, p. 87). They also managed to solve communication problems by negotiating meaning.

Nakatani (2010) studied lower proficiency students in a classroom with a communicative approach to teaching. Once again, students were provided with strategy training and a reference list, but no strategy diary. Based on the data from his study, Nakatani listed the strongest predictors of performance in oral production, the most

significant of which was the use of maintenance strategies. The likely explanation for this result is that engaging the interlocutor by keeping the conversation flowing and creating a more meaningful interaction implies fluency (Nakatani, 2010). Following maintenance strategies was the production rate (i.e., words per minute and length of utterance), the use of signals for negotiation such as “confirmation checks, comprehension checks, and clarification requests during the interaction” (p. 124), and finally the pre-test scores, although this was not a significant predictor of post-test performance. Nakatani found negative correlations with abandonment strategies and the strategy of trying to think in the target language. However, as with pre-test scores, the correlations with performance were not significant. Overall, the use of strategies correlated with higher post-test scores.

Lam (2010) also looked at explicit strategy instruction, comparing low- and high-proficiency students in a high school English class in Hong Kong. Students were taught eight oral communication strategies designed to assist them before, during, and after speech acts, including how to manage the interlocutor’s response. Strategy training was most beneficial to low-proficiency students in terms of vocabulary use, pronunciation, and grammar. Additionally, the low-proficiency students reported using a greater number of strategies than the high-proficiency students. Lam suggests that strategy training may be more beneficial to low-proficiency students on account of their need for strategies to help them through tasks on which they are “linguistically (and perhaps cognitively) weaker” (p. 24). Lam also cites as a possibility Grenfell and Macaro’s (2007) suggestion that high-proficiency students may have already developed a range of strategies. In addition, there may have been too few strategy training sessions and too little practice time to influence the high-proficiency students (Lam, 2010).

Other studies observed student strategy use rather than the effects of strategy training. Chou (2011) looked into the influence of learner strategies on two types of oral

presentation. The students in her study were undergraduate students majoring in French taking a professional English course in Taiwan. Chou hypothesizes that language use strategies, as opposed to learning and metacognitive strategies, were likely most useful for oral presentations, whether they were performed in a group or individually. The longer individual presentations had better structure, a factor related to the use of metacognitive strategies. Students made use of more communication strategies in the group presentation, which allowed them to interact more effectively with the audience, aided their fluency, and helped them avoid simply memorizing text. In both presentations, students made use of cognitive strategies as well as retrieval and rehearsal strategies, although these were used in a different manner. To summarize her findings, group presentations may push students to use more effective communication strategies and thus increase fluency, and individual tasks may push students to use more metacognitive strategies, which lead to better presentation structure.

In her study of native speakers of Chinese and Japanese learning English in Canada, Huang (2010) also found that students used different strategies for different tasks. Rather than explicitly teaching strategies or training students to use them, the teaching methodology in this study aimed to raise students' awareness of their strategy use. Various methods of encouraging reflection, including spoken, written, and group tasks, helped students toward this goal. Huang advises teachers to encourage reflection to help students learn which strategies are most effective for the different tasks they will encounter.

Conclusion

Huang (2010) contrasts the awareness-raising methodology in her study with that of explicit strategy instruction, which does not necessarily encourage “meaningful

evaluation of the most effective strategies for completing the communicative tasks at hand, which learners themselves develop through socially interactive and self-reflective activities” (p. 256). However, the success of the explicit strategy instruction in Lam (2009), Lam (2010), Nakatani (2005), and Nakatani (2010) serves as a testament to its utility, at least in certain contexts. Khan and Victori (2011) offer a different pedagogical tool: a strategy questionnaire as a follow-up to oral communication and a starting point for discussions and assessments of various strategies. They also suggest integrating strategy instruction for the long-term due to its positive effects for oral production. In summary, there are many effective methods for encouraging strategy development in students. Deciding how to teach strategies and determining which ones to focus on will likely depend on teaching context and the needs of individual learners in the classroom.

Summary of Pedagogical Implications

- Not all strategies are equally effective for all tasks and all students. Teachers should encourage students to reflect on which strategies are most effective for different tasks.
- Students can fill out a strategy questionnaire as an introduction to more long-term explicit strategy instruction. Questionnaires are particularly helpful for lower-level students.

FORM INSTRUCTION

Ellis (2002) reviewed form-focused instruction (FFI) research from the 1990s and early 2000s, but drew few firm conclusions on how to teach linguistic forms, largely due to the small number of studies (eleven in total). For free oral production tasks, most studies found increased accuracy after FFI. In studies featuring older learners, however,

FFI was significantly less effective. Ellis found no studies featuring beginning language learners. Ellis found that the effectiveness of FFI appears to depend on the target structure, with less complex structures yielding the best results for oral production. Furthermore, longer treatments of FFI featuring more tasks for students to complete appear to be beneficial for long-term gains in oral production.

Day and Shapson (2001) observed the effects of task-based and FFI of the conditional for grade 6 and 7 French immersion students. They designed instructional materials that would encourage students to use the conditional out of communicative need. Additionally, games and exercises helped reinforce the form, and group- and self-evaluation procedures encouraged students to notice their language use. Day and Shapson found gains for more accurate oral output, though not to the same degree as for written output, perhaps because students tended to avoid using the difficult conditional form when speaking. The authors propose more exposure to a form over several years to promote long-term gains in oral output.

Norris and Ortega (2000) conducted a meta-analysis of 49 FFI studies from 1980 to 1998. Unfortunately, it was impossible for the authors to gauge the effectiveness of different L2 instruction types due to a diversity of research, or to rephrase it more negatively, a lack of focus in the research. Additionally, they found too much variation in interpretations of the different methodologies and flaws in research design, with only 18 percent of studies operationalizing proper control conditions in their opinion. However, there was enough data to suggest that explicit FFI may be more effective than implicit FFI, with the caveat that “most primary research has operationalized implicit treatments in relatively restricted ways, whereas explicit treatments often involve combinations of several instructional components” (p. 483). The data also suggest that Focus on Form (FonF) and Focus on Forms (FonFS) are equally effective for language output quality.

Studies of long-term gains in student production were rare, although when researchers did test for durability of gains, positive results tended to remain, albeit in a somewhat diminished capacity.

A decade later, Serrano (2011) writes that FonFS has not been successful. Although she concedes that the issue is far from settled, she also claims that programs focusing on meaning have been most successful for fluency in oral production, yet students lack grammatical and pragmatic accuracy, a claim also made by much of the research on students of French immersion in Canada (Bouffard & Sarkar, 2008; Day & Shapson, 2001; Lazaruk, 2007) Serrano's study features Spanish partial English immersion grade 6 students and explicit FFI on possessive determiners. The study results demonstrate slight positive effects of the treatment for students' metalinguistic knowledge, performance in metalinguistic tasks, as well as potentially significant effects on oral production tasks, suggesting that FFI may indeed prove beneficial for students in meaning-focused classrooms. Unfortunately the positive effects on oral production tasks seen in this study are tempered by the fact that the control group had very different pre-test scores than the treatment group.

Serrano's (2011) study replicates many aspects of White and Ranta's (2002) study of English immersion students in Quebec. In White and Ranta's study the treatment was much lengthier, and the students were much less accustomed to explicit instruction than the Spanish students. Serrano posits that this novelty seems to have "contributed to raising their interest and motivation" (p. 13). The results for White and Ranta (2002) were much stronger, with students in the treatment group outperforming the control group on measures of oral accuracy and metalinguistic knowledge for possessive determiners. In both of these studies, learner differences were an important factor

affecting output. For example, some students in the control group of White and Ranta's study performed as accurately as the best students in the treatment group.

Bouffard and Sarkar (2008) also explored metalinguistic knowledge by training grade 3 French immersion students in metalinguistic analysis. The students' ability to notice, discuss, and repair errors using metalinguistic terminology significantly improved over the course of the three month study. They were often able to explain the reason for their errors. Their increased ability applied to lexical errors, grammatical errors, and errors related to L1 transfer. Bouffard and Sarkar concede that their research does not necessarily demonstrate a causal link between metalanguage development and L2 development, only that metalanguage assists in students' awareness of their output. The authors conclude by hypothesizing that student awareness might lead to "a higher level of cognitive flexibility" (p. 22). Given that recent immersion research has demonstrated a causal link between attention to form and output under certain circumstances (Day & Shapson, 2001; Lyster, 2004) the pedagogical implications may be more significant than the authors acknowledge.

Conclusion

Due to the issues mentioned in Ellis' (2002) literature review and Norris and Ortega's (2000) meta-analysis, firm conclusions and pedagogical implications from the research remain impossible. Researchers generally agree that form needs to be addressed in the classroom because, as Serrano (2011) argues, a purely meaning-focused classroom can lead to gaps in students' grammatical and pragmatic accuracy. The debate on explicit versus implicit form instruction carries on, as does the debate on FonF versus FonFS. In any case, helping students gain better awareness of form appears to be beneficial for oral production, as does some form of FFI.

Summary of Pedagogical Implications

- FFI can be effective, though possibly less so for older students, and less so for complex structures.
- Teachers should implement tasks in which students must use the target form out of communicative need.
- To promote acquisition, long-term instruction of a form over several years may be necessary.
- Explicit FFI may be more beneficial than implicit FFI.
- Training students in metalinguistic analysis of errors may be beneficial.

CORRECTIVE FEEDBACK

Teachers often use corrective feedback to address their students' non target-like utterances. Feedback can vary in terms of explicitness or implicitness, the degree to which students notice feedback, the degree to which uptake and acquisition occurs, the degree to which students are expected to correct their error, and the degree to which it provides positive or negative evidence. It can also include negotiation of form or meaning. Corrective feedback qualifies as FonF, as it requires "attention to linguistic forms within the context of performing communicative activities" (Ellis, Basturkmen, & Loewen, 2001, p. 281). Therefore, corrective feedback is compatible with teaching methodologies where the primary focus is on meaning. Furthermore, teachers can incorporate corrective feedback into classroom activities without impeding flow (Ellis et al., 2001).

According to Lyster and Saito (2010), recent SLA research concludes that corrective feedback plays a role in developing interlanguage. The general consensus is that FonF through corrective feedback has many benefits and makes a significant impact

on L2 learners' performance with a medium yet substantial effect," however, how to implement it in the classroom is somewhat less clear (Lyster & Saito, 2010, p. 289). Researchers have looked at what types of feedback are most likely to result in student uptake and ultimately result in more accurate output. Though many of the articles reviewed here cover a variety of methods for corrective feedback, recasts have garnered the most interest. This may be due to their high frequency in many contexts (Sheen, 2004).

Nicholas, Lightbown, and Spada (2001) reviewed literature on recasts and concluded that they were generally inadequate for affecting output accuracy, except in the case of phonological errors, a less frequent source of error. In the classroom context, it may be difficult for students to recognize recasts as feedback on form, and it may also be unclear which features of the non target-like utterance were problematic. While the authors acknowledge that the literature includes several definitions of recasts, theirs does not include explicit feedback.

Erlam and Loewen (2010) discuss the difference between implicit and explicit recasts in their study of American university students learning French. They explain that implicit recasts involve a single recast using interrogative intonation, while explicit recasts involve a single repetition of the incorrect utterance with interrogative intonation followed by the corrected form with declarative intonation. They found that there were few differences between the two forms in terms oral output on noun-adjective agreement errors. Regardless of feedback type, students receiving intensive feedback during task-based interactions eventually produced more accurate speech.

Sheen (2004) observed recasts across four contexts and found variance in student uptake and repair. The students in Canadian French immersion and Canadian ESL contexts produced less accurate speech after recasts than did students in Korean EFL and

New Zealand ESL settings. Sheen attributes the differences to the fact that the recasts in both Canadian contexts were more ambiguous and implicit, and failed to provide an opportunity for uptake. In other words, the manner in which teachers provide recasts matters a great deal if it is to eventually affect student output.

Many researchers have sought to compare recasts to other feedback techniques such as prompts. Prompts include “elicitation, metalinguistic clues, clarification requests, and repetition” (Lyster & Saito, 2010, p. 268). These techniques only provide negative evidence and push students to determine the correct form. Lyster (2004) compared recasts and prompts in a grade 5 French immersion classroom. On oral measures, there were few differences between the two groups after a treatment focusing on grammatical gender, and few differences with a third group wherein teachers made use of no particular feedback type. A control group that did not receive FFI on grammatical gender produced significantly less accurate language than the other three groups. Despite the similar results for recasts and prompts, Lyster suggests that recasts are less effective because of the ambiguity they create. Prompts, on the other hand, can assist students’ development of explicit knowledge of grammar rules.

Lyster and Yang (2010) found an advantage for prompts over recasts in their study looking into the effects of corrective feedback on past tense forms in English. The participants in their study were Chinese university students of English. In immediate and delayed post-tests, student usage of regular past tense forms was more accurate for the prompt group than for the recast group, but both groups developed better accuracy of irregular past tense forms. The authors attribute the greater success of prompts to their greater saliency.

Lyster and Mori (2006) looked at recasts, prompts, and explicit correction in two different contexts: Japanese immersion (JI) in the U.S. and French immersion (FI) in

Canada. They found that teachers made use of recasts for various reasons, including scaffolding and positive or negative evidence. The utility of the recasts probably explains their frequency. The authors found different results for uptake and repair across the two contexts. The JI students were more likely to demonstrate uptake and repair after any feedback than the FI students. Recasts were more likely to result in repair for the JI students while prompts were more likely for the FI students, despite teachers providing approximately the same frequency of each type of error correction. Lyster and Mori offer more emphasis on choral repetition of a target model and accurate oral production in the JI context as an explanation for the difference.

Lyster and Mori (2006) come to an interesting conclusion. Their counterbalance hypothesis attempts to explain their findings:

Instructional activities and interactional feedback that act as a counterbalance to the predominant communicative orientation of a given classroom setting will be more facilitative of interlanguage restructuring than instructional activities and interactional feedback that are congruent with the predominant communicative orientation. (p. 294)

Their hypothesis suggests that a form-oriented class like the JI class would benefit from meaning-oriented error correction like recasts. Prompts would be more effective in meaning-oriented classes like the FI class in this study. Although few modern language classes are situated firmly on one end of the spectrum, Lyster and Mori's hypothesis does suggest a logical balance in the classroom.

Lyster (1998) found that the similarity of some teachers' non-corrective repetition and recasts could be a potential cause of recasts' ineffectiveness. When teachers provide recasts in a similar manner to non-corrective repetition, students may not understand the corrective purpose of recasts. In addition, Lyster found that many teachers made use of

positive feedback to express approval of content, yet the form of the utterance may have been incorrect. Again, students may not understand that the positive feedback occurs in relation to content rather than to form.

Lyster and Saito (2010) performed a meta-analysis of the literature on recasts, prompts, and explicit correction. As with recasts, explicit error correction provides positive evidence of the correct form, but it also makes clear the fact that the student used the incorrect form, similarly to Erlam and Loewen's (2010) definition of explicit recasts. The results of Lyster and Saito's meta-analysis suggest that prompts are more effective than recasts overall, largely due to the prompts forcing reformulations, and that "the effects of explicit correction cannot be distinguished from those of recasts or prompts" (p. 283). They found good durability for corrective feedback in delayed post-tests (from two to six weeks) across the literature, a promising finding with implications for acquisition. They also suggest that intensity of corrective feedback on a specific form may be a major factor in the effectiveness of corrective feedback. Age appears to also be a factor, with younger learners benefiting the most, although their treatments tended to be longer than with older learners, which might explain the discrepancy.

Ellis et al. (2001) considered learner uptake after corrective feedback in their study of motivated adult learners in an ESL context. In contrast with most of the studies reviewed here, they found a high level of uptake for recasts. This finding suggests that recasts may yet be effective and should not be discarded for all contexts. The authors found a high level of uptake (72 percent) for all form-focused episodes (FFE) they observed. They believe that a "possible explanation may lie in the fact that the meaning-focused instruction we observed followed a period of focus on forms, which introduced learners to attend generally to form" (p. 311) and enhanced their salience. Student-initiated FFEs were most likely to result in student uptake, while teacher-initiated FFEs

were least likely. The authors suggest that students may be more inclined to pay attention to form when they have noticed their own gaps, and recommend “encouraging students to ask their own questions about form” (p. 314). In terms of feedback type, negotiation of meaning was more successful than negotiation of form, but this may be due to the authors’ coding of lexical items as negotiation of meaning, suggesting that negotiation may be particularly helpful for vocabulary acquisition (Ellis et al., 2001).

In her study of a small sample of 12 middle school Spanish speakers learning English, Alcón-Soler (2009) observed uptake for lexical items and subsequent oral output. As with Ellis et al. (2001), she found that uptake was more frequent in student-initiated FFEs “and when learners provide the information in reactive FFEs than in teacher-initiated or in teacher reactive FFEs” (p. 355). She also finds that recasts are less effective than “elicitation techniques such as questions, metalinguistic clues and clarification questions” (p. 357). Uptake was more likely to occur when teachers provided explicit feedback and when interactions involving negotiation were more complex. Alcón-Soler also discusses the importance of noticing. She found that explicit and implicit feedback techniques both provided opportunities for students to notice form. Noticing led to immediate lexical gains for oral output, however the gains in the delayed post-test were not statistically significant.

Adams, Nuevo, and Egi (2011) observed explicit and implicit corrective feedback in learner-learner interactions rather than between teachers and students. They observed 71 high intermediate adult ESL students in the U.S. engaged in task-based interactions in order to determine if these interactions and the feedback generated within the tasks were successful for the acquisition of English past tense and prepositions of location. They found that in the case of past tense, neither recasts nor explicit corrections seemed to affect learning, although there was a significant correlation between explicit correction of

irregular verbs and their correct usage in an immediate post-test, likely due to the increased salience of irregular verbs. In the case of prepositions of location, the authors found a significant correlation between recasts and the correct use of prepositions of location in a delayed post-test. For all types of feedback provided by students to their peers, much of it was actually not target-like. Students may have been aware of this, rendering any type of learner feedback less convincing than teacher or native speaker feedback. The authors suggest that explicit corrections are more salient to students than recasts, but that overall, there was little gain from learner-learner corrective feedback. It is worth noting that students were not told to focus on the target forms, or even told to focus on accuracy in their interactions. The authors wonder if some students may therefore have been more focused on meaning than on form. They suggest that “in order for learner-learner communicative tasks to promote learning of target forms, teachers may need to direct attention specifically to those forms” (p. 57).

Conclusion

Despite a wealth of pedagogical implications to be gleaned from the research reviewed in this Report, it is difficult to find advice that can be applied to every context. Diane Larsen-Freeman (2003) provides some concrete and concise advice in her guide to teaching grammar. She recommends “affectively supportive, nonjudgmental, judicious, focused feedback that helps students say what they wish to say is vital to successful teaching” (p. 126). Teachers should ensure that they are helping their students develop their interlanguage rather than simply policing it. She also offers that rather than correcting every error, teachers should focus on systematic errors at a point in a students’ language development where they will benefit from feedback, for example when students “need negative evidence in order to eliminate a hypothesis” (p. 133).

Summary of Pedagogical Implications

- Recasts are generally not as effective as prompts except for phonological errors. However, they may be effective if the reason for the recast is clear to the student, if they are explicit, and if students are provided with an opportunity for uptake. They might also be more effective in form-oriented classrooms than in meaning-oriented ones, where prompts are likely to be more effective.
- Intense corrective feedback on a specific form may be more beneficial than sporadic unfocused error correction.
- Teachers should focus on systematic errors and forms appropriate for a student's language level.
- Student-initiated form-focused episodes are more likely to result in uptake than teacher-initiated ones or teacher-reactive ones.
- Teachers should ensure that feedback will not negatively impact students' motivation and confidence and that it helps students to develop their interlanguage.

Discussion

LITERATURE ISSUES

In my introduction I established long-term acquisition of elements that assist with higher quality production as the overarching goal for my students' oral second language skills. However, few of the studies I reviewed performed delayed post-tests to determine long-term benefits, and the ones that did never tested beyond six weeks from the time of the immediate post-test. This raises serious concerns for the effectiveness of any of the teaching methods explored in the literature. Many authors have acknowledged this gap in their research. Samuda and Bygate (2008) cite studies by Mackey (1999) and several articles co-authored by Ellis (Ellis & He, 1999; Ellis & Heimbach, 1997; Ellis, Tanaka & Yamazaki, 1994) that suggest a link between tasks and actual language acquisition. Nonetheless, they explain that "teachers are likely to need more studies like these before they can feel that tasks are a reliable instrument for promoting development" (p. 123). Ellis (2003) believes that "teachers are obliged to assume a relationship between use and acquisition" (p. 138), suggesting that providing situations for students to produce high quality language will result in L2 acquisition. Skehan and Foster (1999) similarly feel that results from immediate post-tests can be extrapolated to longer-term L2 development. It feels to me as though Ellis, and Skehan and Foster are trying to reconcile the gap in research with the reality that it is very difficult to implement delayed post-tests or perform studies over a long period of time. The issue remains somewhat unresolved for me.

Although a thorough analysis of research methodologies is beyond the scope of this Report, authors of meta-analyses of empirical research (Ellis, 2002; Ellis, 2009;

Norris & Ortega, 2000) have pointed out serious methodological flaws in the literature, and most authors acknowledge at least minor flaws in their own research. These are usually due to the difficulty of operationalizing proper control conditions, finding large enough samples of participants, implementing delayed post-tests, or performing studies over several years. Furthermore, measures of complexity, accuracy, and fluency were not operationalized in a uniform manner from study to study. The lack of consistency often renders legitimate comparisons across several studies and contexts difficult.

Samuda and Bygate (2008) point out that “many researchers have doubted the ability or desirability of using tasks to target particular features of language” (p. 122). However, many of the conclusions in the literature are predicated upon this concept. For example, Ellis (2003) provides a list of likely language production outcomes for different types of tasks. He does concede that “particular tasks may predispose learners to engage in certain types of production but they cannot guarantee them” (p. 127). This may be the most realistic manner to describe the potential for certain outcomes based on different forms of input. He is certainly not suggesting that a teacher can plug in a specific task and expect the same outcome in any context.

The disconnect between what is feasible for teachers to implement in classrooms and what researchers have attempted to study raises additional issues. I do not mean to imply that researchers should only measure the results of studies that are replicable by most teachers; however, it is something to take into account when interpreting studies and their results. For example, Yang and Lyster’s (2010) exploration of prompts versus recasts demonstrated the benefits of prompts for regular past tense forms. However, realistically it could take years before a teacher remembered to consistently use prompts instead of recasts for this particular form, especially when recasts remain effective for other errors such as phonological ones. In addition, some studies took place in

laboratories rather than in classrooms, and the results in these cases may not be transferrable to classroom situations.

Three final issues warrant a brief mention. The first stems from the fact that the studies that measured oral production quality rarely measured byproducts such as demotivation as a result of certain tasks, or affective issues like anxiety. Teachers will need to be aware of these possibilities when implementing any tasks or methodologies described in the literature. The second issue relates to individual differences in learners. Few of the studies I reviewed took individual differences into account, although this is hardly surprising given that most of them observed results from groups of students rather than individuals. Learner differences may require teachers to make use of a variety of teaching tools and tasks in the classroom, and ultimately these should be decided upon at the discretion of the teacher. Thirdly, teachers should consider the suitability of different teaching methods and tasks for their unique context, including target language, L1 culture, age of students, and many other individual differences. Most of the studies explored only one educational context, and those that compared results across different contexts found significant variation. This speaks to the importance of considering context when interpreting research results.

CONCLUSION

As I mentioned in the introduction, I became interested in the topic of oral production after researching Canadian French immersion programs, which often emphasize the communication of meaning. Serrano (2011) claims that programs with a focus on meaning have been more successful for fluency than for accuracy in oral production. Many French immersion researchers (Bouffard & Sarkar, 2008; Day &

Shapson, 2001; Lazaruk, 2007) corroborate Serrano's claim. In this Report, I have attempted to find solutions to the issue of non target-like oral production by surveying empirical research on French immersion programs and other L2 instruction settings from the past decade and earlier. I have listed the pedagogical implications for each section, selecting what I believe to be the most useful and effective suggestions for a broad range of educational settings.

I remain convinced of the importance of emphasizing meaning in the classroom. However, I also believe that meaning can be communicated more effectively when students are able to use more accurate and appropriately complex language. Given the success of meaning-focused classrooms for fluency, this model should not be abandoned; rather, it should be built upon by including more FFI. This can be done through a variety of means, and ultimately the most effective teaching methodologies will vary by individual teaching style and educational setting.

There are many possibilities for including more FFI in the meaning-based classroom, many of which are listed in the pedagogical implications for each section in this Report. First, teachers can explicitly teach forms over the course of several years to encourage accuracy as forms become automatized. They can also choose from a variety of techniques for focused intensive corrective feedback on non target-like forms, some more effective for different types of errors than others. Teachers can encourage students to reflect upon different strategies for improving their oral production, and can teach students to notice linguistic forms, lexical items, and useful formulaic sequences. FFI should include many opportunities for practice in meaningful communicative situations.

Throughout my research of the four topics directly related to TBLT (pre-task planning, within-task planning, task repetition, and task design), I have become convinced of the effectiveness of using tasks in the classroom, particularly because they

should be designed to emphasize meaning (Skehan & Foster, 1999). Allowing students time for pre-task and within-task planning can encourage higher quality oral production in terms of accuracy and complexity, although some definitions of TBLT include plenty of practice in time pressure situations as well (Skehan & Foster, 1999). Teachers can also ask students to repeat certain tasks or types of tasks, particularly if the purpose for the repetition extends beyond simply increasing the quality of oral output. Finally, teachers should implement a variety of tasks to maximize the number of linguistic aspects being practiced, and these should be sequenced from least to most complex.

References

- Adams, R., Nuevo, A.M., & Egi, T. (2011). Explicit and implicit feedback, modified output, and SLA: Does explicit and implicit feedback promote learning and learner-learner interactions? *Modern Language Journal*, *95*, 42–63.
- Ahmadian, M. J. (2012). The relationship between working memory capacity and L2 oral performance under task-based careful online planning condition. *TESOL Quarterly*, *46*(1), 165–175.
- Ahmadian, M. J., & Tavakoli, M. (2011). The effects of simultaneous use of careful online planning and task repetition on accuracy, complexity, and fluency in EFL learners' oral production. *Language Teaching Research*, *15*(1), 35–59.
- Alcón-Soler, E. (2009). Focus on form, learner uptake and subsequent lexical gains in learners' oral production. *International Review of Applied Linguistics in Language Teaching*, *47*(3-4), 347–365.
- Aliakbari, M., & Jamalvandi, B. (2010). The impact of “role play” on fostering EFL learners' speaking ability: A task-based approach. *Journal of Pan-Pacific Association of Applied Linguistics*, *14*(1), 15–29.
- Ben Maad, M. R. (2010). Holistic and analytic processing modes in non-native learners' performance of narrative tasks. *System*, *38*(4), 591–602.
- Boers, F., Eyckmans, J., Kappel, J., Stengers, H., & Demecheleer, M. (2006). Formulaic sequences and perceived oral proficiency: Putting a lexical approach to the test. *Language Teaching Research*, *10*(3), 245–261.
- Bouffard, L. A., & Sarkar, M. (2008). Training 8-year-old French immersion students in metalinguistic analysis: An innovation in form-focused pedagogy. *Language Awareness*, *17*(1), 3–24.
- Bygate, M. (1999). Quality of language and purpose of task: Patterns of learners' language on two oral communication tasks. *Language Teaching Research*, *3*(3), 185–214.

- Bygate, M. (2001). Effects of task repetition on the structure and control of language. In M. Bygate, P. Skehan, & M. Swain (Eds.), *Researching pedagogic tasks: Second language learning, teaching and testing* (pp. 23–48). Essex: Longman.
- Bygate, M., Skehan, P., & Swain, M. (2001). *Researching pedagogic tasks: Second language learning, teaching and testing*, (Eds.). Essex: Longman.
- Chou, M. (2011). The influence of learner strategies on oral presentations: A comparison between group and individual performance. *English for Specific Purposes*, 30(4), 272–285.
- Day, E. M., & Shapson, S. M. (2001). Integrating formal and functional approaches to language teaching in French immersion: An experimental study. *Language Learning*, 51, 47–80.
- Elder, C., & Iwashita, N. (2005). Planning for test performance: Does it make a difference? In R. Ellis (Ed.), *Planning and task performance in a second language* (pp. 220–238). Amsterdam: John Benjamins Publishing Company.
- Ellis, R. (2002). Does form-focused instruction affect the acquisition of implicit knowledge? A review of the research. *Studies in Second Language Acquisition*, 24(2), 223–36.
- Ellis, R. (2003). *Task-based language learning and teaching*. Oxford: Oxford University Press.
- Ellis, R. (2005). *Planning And task performance in a second language*, (Ed.). Amsterdam: John Benjamins Publishing Company.
- Ellis, R. (2009). The differential effects of three types of task planning on the fluency, complexity, and accuracy in L2 oral production. *Applied Linguistics*, 30(4), 474–509.
- Ellis, R., Basturkmen, H., & Loewen, S. (2001). Learner uptake in communicative ESL lessons. *Language Learning*, 51(2), 281–318.
- Ellis, R., & He, X. (1999). The role of modified input an output in the incidental acquisition of word meanings. *Studies in Second Language Acquisition*, 21(2), 285–301.
- Ellis, R., & Heimbach, R. (1997). Bugs and birds: Children’s acquisition of second language vocabulary through interaction. *System*, 25(2), 247–259.

- Ellis, R., Tanaka, Y., & Yamazaki, A. (1994). Classroom interaction, comprehension, and the acquisition of L2 word meanings. *Language Learning*, 44(3), 449–491.
- Ellis, R., & Yuan, F. (2005). The effects of careful within-task planning on oral and written task performance. In R. Ellis (Ed.), *Planning and task performance in a second language* (pp. 167–192). Amsterdam: John Benjamins Publishing Company.
- Erlam, R., & Loewen, S. (2010). Implicit and explicit recasts in L2 oral French interaction. *Canadian Modern Language Review*, 66(6), 877–905.
- Foster, P., & Skehan, P. (1999). The influence of source of planning and focus on task-based performance. *Language Teaching Research*, 3(3), 215–47.
- Gatbonton, E., & Segalowitz, N. (2005). Rethinking communicative language teaching: A focus on access to fluency. *Canadian Modern Language Review*, 61(3), 325–353.
- Gilabert, R. (2007). Effects of manipulating task complexity on self-repairs during L2 oral production. *International Review of Applied Linguistics in Language Teaching*, 45(3), 215–240.
- Gilabert, R., Baron, J., & Llanes, A. (2009). Manipulating cognitive complexity across task types and its impact on learners' interaction during oral performance. *International Review of Applied Linguistics in Language Teaching*, 47(3-4), 367–395.
- Grenfell, M., & Macaro, E. (2007). Claims and critiques. In A.D. Cohen and E. Macaro (Eds.), *Language learner strategies* (pp. 9–28). Oxford: Oxford University Press.
- Huang, L.-S. (2010). Do different modalities of reflection matter? An exploration of adult second-language learners' reported strategy use and oral language production. *System*, 38(2), 245–261.
- Kawauchi, C. (2005). The effects of strategic planning on the oral narratives of learners with low and high intermediate L2 proficiency. In R. Ellis (Ed.), *Planning and task performance in a second language* (pp. 143–164). Amsterdam: John Benjamins Publishing Company.
- Khan, S., & Victori, M. (2011). Perceived vs. actual strategy use across three oral communication tasks. *International Review of Applied Linguistics in Language Teaching*, 49(1), 27–53.

- Lam, W. Y. K. (2009). Examining the effects of metacognitive strategy instruction on ESL group discussions: A synthesis of approaches. *Language Teaching Research*, 13(2), 129–150.
- Lam, W. Y. K. (2010). Implementing communication strategy instruction in the ESL oral classroom: What do low-proficiency learners tell us? *TESL Canada Journal*, 27(2), 11.
- Larsen-Freeman, D. (2003). *Teaching language: From grammar to grammaring*. Boston: Heinle-Thomson.
- Lazaruk, W. (2007). Linguistic, academic, and cognitive benefits of French immersion. *Canadian Modern Language Review*, 63(5), 605–627.
- Lynch, T. (2007). Learning from the transcripts of an oral communication task. *ELT Journal*, 61(4), 311–320.
- Lynch, T., & Maclean, J. (2001). “A case of exercising”: Effects of immediate task repetition on learners’ performance. In M. Bygate, P. Skehan, & M. Swain (Eds.), *Researching pedagogic tasks: Second language learning, teaching and testing* (pp. 141–162). Essex: Longman.
- Lyster, R. (1998). Recasts, repetition, and ambiguity in L2 classroom discourse. *Studies in Second Language Acquisition*, 20(1), 51–81.
- Lyster, R. (2004). Differential effects of prompts and recasts in form-focused instruction. *Studies in Second Language Acquisition*, 26(3), 399–432.
- Lyster, R., & Mori, H. (2006). Interactional feedback and instructional counterbalance. *Studies in Second Language Acquisition*, 28(2), 269–300.
- Lyster, R., & Saito, K. (2010). Oral feedback in classroom SLA: A meta-analysis. *Studies in Second Language Acquisition*, 32(2), 265–302.
- Mackey, A. (1999). Input, interaction, and second language development: An empirical study of question formation in ESL. *Studies in Second Language Acquisition*, 21(4), 557–87.
- Mennim, P. (2003). Rehearsed oral L2 output and reactive focus on form. *ELT Journal*, 57(2), 130–38.

- Nakatani, Y. (2005). The effects of awareness-raising training on oral communication strategy use. *Modern Language Journal*, 89(1), 76–91.
- Nakatani, Y. (2010). Identifying strategies that facilitate EFL learners' oral communication: A classroom study using multiple data collection procedures. *Modern Language Journal*, 94(1), 116–136.
- Nicholas, H., Lightbown, P. M., & Spada, N. (2001). Recasts as feedback to language learners. *Language Learning*, 51(4), 719–58.
- Norris, J. M., & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning*, 50(3), 417–528.
- Robinson, P. (2001). Task complexity, task difficulty, and task production: Exploring interactions in a componential framework. *Applied Linguistics*, 22(1), 27–57.
- Samuda, V., & Bygate, M. (2008). *Tasks in second language learning*. Basingstoke: Palgrave Macmillan.
- Serrano, R. (2011). From metalinguistic instruction to metalinguistic knowledge, and from metalinguistic knowledge to performance in error correction and oral production tasks. *Language Awareness*, 20(1), 1–16.
- Sheen, Y. (2004). Corrective feedback and learner uptake in communicative classrooms across instructional settings. *Language Teaching Research*, 8(3), 263–300.
- Skehan, P., & Foster, P. (1999). The influence of task structure and processing conditions on narrative retellings. *Language Learning*, 49(1), 93–120.
- Stengers, H., Boers, F., Housen, A., & Eyckmans, J. (2011). Formulaic sequences and L2 oral proficiency: Does the type of target language influence the association? *International Review of Applied Linguistics in Language Teaching*, 49(4), 321–343.
- Taguchi, N. (2007). Task difficulty in oral speech act production. *Applied Linguistics*, 28(1), 113–135.
- Taguchi, N. (2008). Building language blocks in L2 Japanese: Chunk learning and the development of complexity and fluency in spoken production. *Foreign Language Annals*, 41(1), 132–156.

- Tavakoli, P., & Foster, P. (2008). Task design and second language performance: The effect of narrative type on learner output. *Language Learning*, 58(2), 439–473.
- Tavakoli, P., & Skehan, P. (2005). Strategic planning, task structure, and performance testing. In R. Ellis (Ed.), *Planning and task performance in a second language* (pp. 239–273). Amsterdam: John Benjamins Publishing Company.
- White, J., & Ranta, L. (2002). Examining the interface between metalinguistic task performance and oral production in a second language. *Language Awareness*, 11(4), 259–90.
- Wigglesworth, G. (2001). Influences on performance in task-based oral assessments. In M. Bygate, P. Skehan, & M. Swain (Eds.), *Researching pedagogic tasks: Second language learning, teaching and testing* (pp. 186–209). Essex: Longman.
- Yang, Y., & Lyster, R. (2010). Effects of form-focused practice and feedback on Chinese EFL learners' acquisition of regular and irregular past tense forms. *Studies in Second Language Acquisition*, 32(2), 235–263.
- Yuan, F., & Ellis, R. (2003). The effects of pre-task planning and on-line planning on fluency, complexity, and accuracy in L2 monologic oral production. *Applied Linguistics*, 24(1), 1–27.