Motivation, language development, and lesson design in task-based lessons in two foreign language classrooms

Renee Waara
Høgskolen i Nesna

The fieldwork for this study is conducted by two PPU students. Lessons are designed and implemented following principles of task-based learning in upper secondary school, one in German and one in Spanish. A pre- and post-survey is collected that maps affective variables and preferred types of language activities. Spoken pre- and post-tests are collected by the students using wordless comic strips. Transcripts are analyzed using standardized measures for fluency, lexical complexity, structural complexity, and accuracy. Language production in the two classes improved in different areas after a 16 lesson project teaching. Correlations between lesson design, language production differences, and motivational factors are identified and discussed in light of the Cognition Hypothesis.

Introduction

In a previous study (Waara, 2009), teachers are aware of the value of encouraging planning time and promote it actively, and yet these lessons are unable to accumulatively promote neither structural complexity nor accuracy in a post-test in the project group. A tentative explanation is suggested that input needed more emphasis.

The attentional demands of the narrative task are distributed differently between the project group and the control group. The 5th graders who are not exposed to the project teaching are very careful and very accurate in their picture narratives and improve in accuracy, but do not improve in structural complexity, in contrast to the experiment group who improve in fluency and lexical complexity. This is predominantly in line with Foster and Skehan (1996) who predict that attentional resources are limited and in competition such that either structural complexity will prosper or accuracy, but rarely both.

The current study emerges from the previous study. Input and pushing language production in speech are emphasized, in a context that seeks to explore what the classroom setting yields in terms of interaction between lesson design and language learning processes. The focus in this study is on language input and output, in a task-based learning framework that emphasizes noticing and use of the target language in the classroom, in other words, components that should contribute to learner language development.

Study background

The notion of 'attention' in language learning has received considerable amount of focus in the past two decades. Schmidt (1995; 2001) claims that 'noticing' or conscious attention is necessary for language learning to take place. However, working memory is limited and given that learner language is only partially automatic we assume that language production requires some form of attention (de Bot, 1992). Studies on the production mechanisms involved in task-based language learning have looked at pre-task and on-line planning conditions (Mehnert, 1998; Ortega, 1999; Yuan & Ellis, 2003). By adjusting planning time the learner's attention is directed to specific dimensions of language production such as fluency, complexity or accuracy.

In Robinson's Cognition Hypothesis (2001a; 2001b; 2002; 2003) he proposes that pedagogical tasks be sequenced based on cognitive complexity. Robinson's hypothesis makes several predictions, one of the predictions is that by increasing the cognitive demands of tasks the learner will be pushed to «greater accuracy and complexity of L2 production in order to meet the consequently greater functional/communicate demands they place on the learner» (2003, p. 45).

Robinson (2003) distinguishes between two categories of the dimensions of task complexity, *resource-dispersing* dimension and *resource-directing* dimension. The resource-dispersing dimension does not provide learners with any bits of language or 'language code' to complete the task, rather it refers to planning time, the status of prior knowledge, the number of tasks that have to be carried out simultaneously, in other words, how attentional resources are dispersed. This dimension does not develop language, but allows access to already existing L2 knowledge. The resource-directing dimension provides learners with a need for specific aspects of language. The demands of the task are cognitively complex and require the learner to discover the language they need to complete the task and consequently promotes greater syntacticization and grammaticization of current interlanguage. This dimension contributes to the development of learner language.

In the 5th grade study, resource-dispersing, i.e. planning time, is emphasized, whereas in this study, the resource-directing dimension, i.e. the language bits, are in focus.

Study questions

Very few studies in this field are conducted in the classroom. Experimental studies are typically conducted on university students or other groups of adult learners. Isolating aspects of tasks in relation to language production and language development in experimental settings sheds light on how task complexity affects language development and provides systematic ways of organizing syllabi and designing lessons. For example, we know that attentional resources are limited and that this affects learner performance in trade-off effects, such that a learner does not have resources to attend to all dimensions of language production, i.e. fluency, lexical complexity, structural complexity, and accuracy. But we do not know how this is realized in the classroom and whether or not other elements of interaction, e.g. related to lesson design, role of teacher, teacher attitudes, and level of development, will have an influence on the trade-off effects.

Therefore it is interesting to apply these ideas and principles to classroom use and perhaps extend our practical knowledge about consequences of lesson design and learner language development in the classroom. There are three main questions in this study.

- Does a change in motivation correlate with improvement? In other words, are the pupils who improve most in the areas of language production measured, more motivated after the project?
- Does language production improve after the 16 lesson teaching project in areas of fluency, lexical complexity, structural complexity, and accuracy?
- Are there systematic differences in lesson design in terms of resource-directing dimensions of complexity between the German and Spanish classes?

Project design and material

The task as the basic unit of lesson design can be classified based on features of complexity. Thus, task-based learning is chosen because using it allows for a comparison between task complexity and language improvement. Lessons are designed using basic principles of task-based learning, i.e. pre-task phase, task cycle that emulates real world communication, preparation/planning phase, and performance as discussed in Willis (1996). The language focus component is to be part of the lessons, but this tends to be deleted quite frequently due to lack of time at the end of a lesson. The issue is discussed during the project as a possible source of negative influence on accuracy.

The student teachers are given permission to use their practice teaching classes in the project on the condition that the material on the syllabus is covered. This meant that we are in some cases limited in our topics and material, but we still are able to apply the task-based model. Both classes are in the first year of upper secondary school, with one difference. The German class had German in lower secondary school, while the Spanish class did not. The classes consist of 25 pupils, ages 16–17.

The pupils are recorded before and after the 16-lesson teaching block, which occurred over a period of approximately 12 weeks. During the recording procedure, pupils are given as much time as they want to prepare their stories, and any questions they have about content or clarification of the pictures is provided before their stories are recorded. The idea is that planning time should not be a hindrance, and that pupils are free to show their learner language production.

The material is transcribed by the student-teachers. Pauses are indicated in the transcripts with periods for complete breaks marked by falling intonation, commas indicate a shorter break and then the speaker continues an utterance, and three dots indicate a perceptible break. These phenomena are used in identifying AS-units (defined below in this section). The comic strips and procedures for data collection and analysis are adapted from Gilabert (2005).

Speech production is measured in terms of fluency, lexical complexity, structural complexity, and accuracy. The calculations are described in detail.

Unpruned speech and pruned speech are used to measure fluency. Unpruned speech is calculated by counting the number of syllables and dividing by the total number of seconds and multiplying by 60. Pruned speech is calculated in the same manner, but excludes repetitions, self-corrections, false starts, and comments in Norwegian. These

two independent measures yielded the same results with the exception of one case in Spanish.

Lexical complexity is measured using three different measures, percentage of lexical words, ratio of lexical to functional words, and Guiraud's index. The percentage of lexical words is calculated by taking the number of lexical words in the transcriptions and dividing it by the total number of words and multiplying it by 100. The ratio of lexical to function words is calculated by taking the number of lexical words and dividing it by the number of function words and multiplying it by 100.

Guiraud's index of lexical richness is calculated by taking the number of types divided by the square root of the total number of words in a text (types/√tokens). Wordsmith, a concordance program, is used to identify types.

Structural complexity is basically the ability to put more than one constituent together in a meaningful way. In other words, when coordination or subordination are put together such that the bits have meaning, then we have complexity. False starts, repetition, corrections, are not counted as AS-units in this calculation.

The measure used for calculating structural complexity is number of clauses divided by the number of Analysis of Speech Units (AS-units). AS-units are described in Foster, Tonkyn & Wigglesworth, (2000, p. 365) as «a single speaker's utterance consisting of *an independent clause, or sub-clausal unit*, together with any *subordinate clause(s)* associated with either». AS-units are chosen because they best are able to deal with features characteristic of spoken data, such as intonation and pauses. A clause is defined as a finite or non-finite verb plus one additional clause element.

The measure used for calculating accuracy is initially error free AS-units. Further measures of accuracy are considered including, ratio of errors to words, ratio of errors to AS-units, and number of errors divided by self-repairs.

Data presentation and discussion

Is motivation related to language improvement?
The motivation survey contains three sections, affective variables, preferred activities and open-ended questions. Only the first two sections are presented in this study, as the open-ended question is directed towards the pupils' experiences of the student-teacher. It was thought that there may be a correlation between motivation for learning a foreign language and language improvement. However, there are no correlations on the individual level between affective variables and measurable language improvement to support this idea. Affective variables turn out to be more interesting when compared to task complexity and lesson design. For example, the difficulty of a task should be relatively high if our goal is to challenge pupils to perform and discover language they need, subsequently leading to language development. The degree of relaxation and frustration, however, is a delicate balance. A high level of relaxation might indicate that there is a lack of challenge and that it does not lead to learner language development. The degree of how interesting pupils perceive the lesson is also central; pupils need to be able to relate, in some way, to the lesson material in an engaging manner.

In the survey that is conducted before and after the project teaching, pupils are asked how they experience their lessons on a 6-step scale from 'very relaxed' to 'very frus-

trated', 'very easy' to 'very difficult', and 'very interesting' to 'very uninteresting'. Both the German class and the Spanish class perceive their classes as being 'more relaxed' after the project teaching, 13 and 14 % respectively. 21 % of the pupils in the Spanish class perceive the teaching as less relaxed after the project and only 2 % of the pupils in the German class. Based on this variable alone, assuming that they are less relaxed because they find the lesson difficult, we might expect to find that the Spanish class improved more than the German, and this is indeed supported by the data in some areas of language production.

The difficulty level does not vary for the German pupils much between the pre- and post-surveys. However, the magnitude is significant. 60 % of the German pupils perceive their lessons as 'a little difficult' in the pre-survey and 57 % in the post-survey, and 20 % as 'difficult' in the pre-survey, and 21.5 % in the post-survey. These two categories combined, i.e. 80 and 78.5 %, suggest that the German pupils associate the notion of 'difficult' to their lessons.

The Spanish class gives a slightly different picture. 57 % of the Spanish pupils perceive their lessons as 'a little difficult' in the pre-survey and 50 % in the post-survey. In contrast to the German pupils, only 7 %, in both the pre- and post-survey, perceive the lessons as 'difficult'.

The German pupils perceive their lessons as more interesting after the project teaching by almost 19 %, whereas the Spanish show virtually no change.

Activity preference is registered, on the one hand, to see if the project teaching that is focused on speaking activities will affect pupils' preference for speaking activities, and on the other hand, to see if there are correlations between language production improvement and activity types. The survey question used to identify activity preference types asks the pupils to rank the following activities: reading, writing, speaking, grammar, and listening. The responses from three pupils from each class who demonstrate most improvement in terms of fluency, lexical, structural and accuracy are compared. The two top responses from each pupil's pre- and post-surveys are compared and consistently one activity from each emerged as the preferred activity. In cases where both activities were ranked in the top two, both activities are included. The same procedure is applied to the least preferred activity.

In a comparison of pupils who improve most in language production, we find that doing grammar activities ranks higher as a preferred activity for German pupils than for Spanish pupils. In contrast, grammar ranks high as least favorite activity for Spanish pupils. Listening activities rank high as the least favorite activity type for the German pupils.

Preferred activities reflect on what pupils perceive themselves as competent at or are comfortable doing. Therefore it is rather odd that the pupils who improve in the most areas of language production do not prefer speaking activities. It is even more difficult to explain why the preferred activity for pupils who improve the least overall is speaking for both the Spanish and German pupils. Grammar and listening activities rank high as least favorite types of activities for both Spanish and German pupils who improve the least overall.

Does language production improve after the 16-lesson teaching project in areas of fluency, lexical complexity, structural complexity and accuracy?

Improvement is a simple calculation in which the post-test results are better than the pre-test results. When two thirds of the group improves, the group is considered to show improvement. These numbers are indicated in bold in Tables 1 and 2. There are 14 subjects, pupils who are tested, in each group. Accuracy is initially measured as error free units, but it did not capture the nuance of the amount of language that is produced and therefore more measures of accuracy are considered and described in Table 2.

Measure	German	Spanish
Fluency	6	**13**
Lexical complexity	**10**	**13**
Structural complexity	7	**14**
Accuracy	**12**	3

Table 1: Measures that indicate improvement in language production between pre- and post-tests, N=14

Measure	German	Spanish
Error free AS-units	**12**	3
Ratio of errors to words	**10**	**10**
Ratio of errors to AS-units	**11**	7
Number of errors/self-repairs	3	4

Table 2: Accuracy measures that indicate improvement in language production between pre- and post-tests, N=14

As shown in Table 1, the two classes improve in different areas of language production. The German class improves in accuracy and lexical complexity, but not in fluency and structural complexity. The Spanish class improves in all areas except accuracy, with the exception of ratio of errors to words as illustrated in Table 2. The differences found in improvement are attributed to the project teaching. The German class is very accurate but produces overall less language than the Spanish class.

Errors are defined in terms of lexical, morphologic, and syntactic anomalies that are normally not found in the language or produced by a native speaker. Self-repairs are occurrences where pupils monitor and adjust their speech accordingly. There are more occurrences of self-repairs in the Spanish data, which suggests that they monitor their language and have the appropriate knowledge to do it. The German class is better in terms of error free AS-units, but is conservative in their speech production as is reflected in the ratio of errors to words.

Do differences in lesson design between German and Spanish account for differences in language improvement?

The PPU students describe their lessons in detail in the projects. Using the students' project reports in combination with the discussions that we had after each lesson, lessons are categorized using Robinson's divisions for resource-directing and resource-dispersing dimensions of complexity. Robinson (2003) proposes that:

> Increasing complexity along resource-directing dimensions can be expected to lead the learner to attempt to map the conceptual/functional requirements of tasks onto speech, in such a way as to affect fluency negatively, but, in selected domains, to facilitate the development of increased accuracy and complexity of production [...]. In contrast, increasing complexity along resource-dispersing dimensions can be expected to affect fluency, as well as accuracy and complexity, negatively, since it creates problems for learners attempting to access their current repertoire of L2 knowledge (p. 59).

The developmental and the performative dimensions of complexity can be grouped into four combinations:

- low performative and low developmental complexity
- high performative and low developmental complexity
- low performative and high developmental complexity
- high performative and high developmental complexity

The low performative dimension refers to the presence of time for planning utterances, the availability of prior knowledge, and the singularity of task. The high performative dimension refers to lack of time for planning, no prior knowledge that is necessary to solve the task, and the duality of task. Low developmental complexity refers to a low number of elements that are needed to conduct the task, no reasoning is required, and the task takes place in the 'here-and-now'. High developmental complexity refers to several elements that are needed to conduct the task, reasoning is required, and the task takes place in the 'there-and-then'.

Robinson applies these combinations to single lessons in experimental settings. If we apply the principles of these dimensions to an extended teaching block to which we have access to the degree of complexity of each lesson, what do we find?

German increases in accuracy so, according to the Cognition Hypothesis, we might expect to find that the 16 lessons have a high degree of complexity. This turns out to be the case and the data supports this idea, whereas the Spanish class gives us a slightly more complex picture. The class improves in all areas except accuracy, although they do improve in terms of ratio of errors to words. Taking into consideration the level of task complexity and the areas of improved language production presents us with a complicated picture in relation to the predictions made by the Cognition Hypothesis. Accuracy is a measure of language development, but this does not appear to improve as a result of the project teaching, whereas the lexical complexity and structural complexity measures are also measures that indicate language development has occurred.

The level of complexity of each lesson is evaluated in terms of resource-directing dimension (number of elements, here-and-now vs. there-and-then, reasoning) and resource-dispersing (planning, prior knowledge, and number of tasks) as described above. In cases where the lesson/task contained combinations of categories, these are indicated with backslashes.

Lessons	1	2	3	4	5	6	7	8
German	3	3	1	2	4	3	3/2	3/4
Spanish	1	1	-	4	1/2	3	1	4

Table 3: Level of task complexity per lesson

According to the mode there is an identifiable difference between the types of lessons in German and Spanish. German has a mode of 3 and Spanish has a mode of 1 leaning towards 2. In other words, the German lessons more frequently contain several elements, require reasoning, occur in the 'there-and-then', and at the same time allow for time for planning, use prior knowledge and are single task based. The Spanish lessons more frequently contain few elements, do not require reasoning, occur in the 'here-and-now', and allow time for planning, use prior knowledge and are single task based.

German lessons are generally more complex than the Spanish lessons, and support Robinson's predictions in that fluency is affected negatively and accuracy positively. However, the Spanish lessons are generally less complex and the pupils improve in every language dimension except accuracy, and this does not follow Robinson's prediction that task complexity drives language development forward and causes learner language development.

Concluding remarks

The focus in this study is on maximizing input and stimulating output. Over a period of 16 lessons applying a specific method in a careful lesson-designing process, we wanted to see if we could influence the language production of these learners. The results diverge in systematic ways between the two classes. The German class, on the one hand, improves in accuracy and lexical complexity. The Spanish class, on the other hand, improves in fluency, lexical and structural complexity, but not in over-all accuracy.

The conditions for the lesson design process are similar for the two classes in that certain restraints for using the textbook topics are imposed and that both student-teachers are native speakers in their respective languages. This is considered an advantage as it is instrumental in achieving the goal of providing ample input.

Even though the lesson design process is similar, a detailed analysis of each lesson in terms of the resource-directing and resource-dispersing dimensions reveals that the lessons are not similar in the degree of complexity. While both classes are exposed to the low levels of the performative dimension (i.e. the presence of time for planning utterances, the availability of prior knowledge, and the singularity of task), the German class more frequently experiences lessons with high developmental complexity (i.e. several elements are needed to conduct the task, reasoning is required, and the task takes place in the 'there-and-then'). The Spanish class more frequently experiences lessons with low developmental complexity (i.e. a low number of elements are needed to conduct the task, no reasoning is required, and the task takes place in the 'here-and-now'). These observations coincide with the pupils' own perceptions of the lessons in that the German pupils, overall, perceived the lessons as more difficult than the Spanish pupils.

The Spanish class is less relaxed after the project by 21 % and their language production improves in everything except accuracy as measured by error free AS-units. And yet the Spanish class is relaxed enough to produce extensive language and they produce more language in the post-test with respectively fewer overall errors than the German class. A partial explanation could be that the Spanish class is driven to produce language and is not concerned with form, i.e. accuracy, and that this is related to their attitudes toward grammar, which is identified as their least favorite activity.

The German class does not perceivably experience lessons as more difficult after the project, but they do experience the lessons as more interesting by almost 19 %. Maybe the 'interesting' variable is just as or more important for language development as is the level of task complexity. Accuracy as a measure of language development may also be affected by the way grammar activities are perceived by the German class, in that there is a strong preference for grammar activities.

References

de Bot, K. (1992). A bilingual production model: Levelt's «speaking» model adapted. *Applied Linguistics, 13*(1), 1–24.

Foster, P., & Skehan, P. (1996). The influence of planning and task type on second language performance. *Studies in Second Language Acquisition, 18*(3), 299–323.

Foster, P., Tonkyn, A., & Wigglesworth, G. (2000). Measuring spoken discourse: A unit for all reasons. *Applied Linguistics, 21* (3), 333–353.

Gilabert, R. (2005). *Task complexity and L2 narrative oral production.* Unpublished dissertation. University of Barcelona, Spain.

Mehnert, U. (1998). The effects of different lengths of time for planning on second language performance. *Studies in Second Language Acquisition, 20*(1), 83-108.

Ortega, L. (1999). Planning and focus on form in L2 Oral Performance. *Studies in Second Language Acquisition, 21*(1), 109–148.

Robinson, P. (2001a). Task complexity, cognitive resources and syllabus design: A triadic framework for examining task influences on SLA. In P. Robinson (ed.), *Cognition and second language instruction* (p. 185–316). New York: Cambridge University press.

Robinson, P. (2001b). Task complexity, task difficulty and task production: Exploring interactions in a componential framework. *Applied Linguistics, 22*(1), 27–57.

Robinson, P. (2002). Individual differences in intelligence, aptitude and working memory during adult incidental second language learning: A replication and extension of Reber, Walkendfield and Hernstadt (1991). In P. Robinson (ed.), *Individual differences and instructed language learning* (p. 211–266). Amsterdam: Benjamins.

Robinson, P. (2003). The cognition hypothesis, task design, and adult task-based language learning. *Second Language Studies, 21* (2), 45–105.

Schmidt, R (1995). Consciousness and foreign language learning: A tutorial on the role of attention and awareness in learning. In R. Schmidt (ed.), *Attention and awareness in foreign language learning* (p. 1–64). Honolulu, HI: University of Hawaii Press.

Schmidt, R. (2001). Attention. In P. Robinson (ed.), *Cognition and second language instruction* (p. 1–33). New York: Cambridge University Press.

Waara, R. (2009). Connecting pupils, students and theory in a project: Planning time in the classroom, a crucial element in lesson design. In B. Groven, T. M. Guldal, O. F. Lillemyr, N. Naastad, & F. Rønning (eds.), *FoU i Praksis 2008. Rapport fra konferanse*

om praksisrettet FoU i lærerutdanningen (p. 317–326). Trondheim: Tapir Akademisk Forlag.

Willis, J. (1996). *A framework for task-based learning*. London: Longman.

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.