

Designing a communicative foreign language assessment program for Dutch secondary schools: A design-based research project within a professional learning community (PLC)

Charline Rouffet

HU University of Applied Sciences Utrecht, the Netherlands

Catherine van Beuningen

HU University of Applied Sciences Utrecht, the Netherlands

Rick de Graaff

HU University of Applied Sciences Utrecht, the Netherlands

Abstract

While Communicative Language Teaching (CLT) is nowadays widely recognized, the implementation of CLT in foreign language (FL) classrooms remains difficult. In the Netherlands, communicative learning goals have been formulated at the national level, but in daily practice assessments and learning activities tend to focus on knowledge of grammar rules and vocabulary out of context.

Under the principle of constructive alignment, assessment and learning activities should be in line with learning objectives in order to enable effective teaching. Evaluation in particular plays a key role, because it has a direct impact on teachers' pedagogical choices and students' learning behavior (i.e., washback effect).

To enhance alignment in the Dutch FL curricula of lower form education, the first author and 21 FL teachers collaborated to create a realistic, theoretically grounded communicative assessment program that would enable positive washback. This paper takes inventory of the challenges faced by teachers during the co-design process and the decisions that have been made to overcome them.

Introduction

Communicative language teaching (CLT) aims to promote the development of communicative language skills in realistic situations. The approach emphasizes the use of language through meaningful spoken and written interaction rather than through memorization and the learning of grammatical rules out of context (Larsen-Freeman & Anderson, 2011). CLT is nowadays widely accepted and recognized as a productive pedagogical approach in foreign language (FL) education. However, the implementation of CLT in daily teaching practices remains difficult (Kissau, Rodgers, & Haudeck, 2015). This situation also seems to occur in the Netherlands. In Dutch secondary education, learning objectives at the national level are formulated from a communicative perspective, but in teaching and assessment practices teachers tend to focus on knowledge of grammar rules, vocabulary and chunks taught out of context (Fasoglio, de Jong, Pennewaard, Trimbos, & Tuin, 2015; West & Verspoor, 2016). As a result, students who perform well during their school career and up to their final exams still have difficulties or are even unable to communicate in the foreign languages they have learned (Schnabel et al., 2016). One of the explanations for this situation could be a lack of constructive alignment in Dutch foreign language curricula.

According to Biggs (1996), the effectiveness of any curriculum depends on its 'constructive alignment': the coherence, or degree of alignment, between learning objectives, learning activities and assessment practices. In the development of an aligned curriculum, Biggs emphasizes the importance of identifying the goals that students need to achieve in assessments before organizing teaching and learning activities, as tests are known to influence both teaching and learning; this influence is known as 'washback' (Alderson & Wall, 1993; Green, 2007). This washback can be positive when tests are aligned with learning goals, but can be negative when alignment is lacking. In the context of FL teaching, where students have to learn to communicate in a FL

(learning objective), this means that assessment and learning activities are relevant (i.e., aligned) when they are communicative in nature. To ensure positive washback (or to prevent its negative counterpart), FL teachers should provide communicative tests in which learners are asked to perform authentic tasks similar to those they encounter in real life (Morrow, 2018).

Brown (2005) identifies five requirements that communicative assessment activities should meet:

1. Meaningful communication: Tasks should be connected to students' experiences and perceptions.
2. Authentic situation: Tasks should be linked to real-life situations.
3. Unpredictable language input: Students should be able to respond to unprepared questions or comments.
4. Free production: Students should be able to show that they can give their own content to a conversation.
5. Integration of language skills: Tasks should encourage students to use different language skills in an integrated way, as it is often the case in real-life communication.

In lower-form secondary education in the Netherlands (students aged 12 to 15), test formats and contents are chosen by departments of individual schools. Foreign language teachers are expected to develop their curriculum and tests in line with the communicative objectives formulated at the national level: students should be able to use different language skills in authentic situations in order to communicate effectively in a FL (College voor Toetsen en Examens, 2017). However, teachers often use tests that come with widely-used textbooks, which tend to focus on assessing vocabulary, grammar, and receptive skills (reading and listening). Productive skills (speaking and writing) are tested less often, and if they are, this is usually done by asking students to translate sentences or reading texts out loud (Fasoglio et al., 2015). Such tests do not adequately introduce real-life unpredictability (Brown, 2005) in the form of unknown or unprepared questions and fail to test communicative competence, i.e. the ability to adapt FL knowledge and skills to new situations.

Considering the washback effect of tests, this could explain the difficulty of implementing communicative teaching in Dutch FL classrooms; if tests are not sufficiently communicative, learning activities will not be sufficiently communicative either. In (partly) decentralized educational systems such as the Dutch one, teachers have a lot of control over teaching materials and assessments which would, in principle, grant them with ample opportunity to generate positive washback (Hakim, 2018). Harding (2014), however, reports that individual teachers face challenges in designing communicative language tests, due to practical concerns and different interpretations of CLT. As a result, assessment tasks and rating scales designed by teachers, if any, are often developed or adapted intuitively and are not always able to validly measure students' communicative competence (Fulcher, 2003). It is therefore important to better understand which practical and conceptual challenges teachers face when designing tests and to identify which decisions they can make to overcome them.

The practical objective of the current design study (part of a larger designed based-research project) was to, collectively with FL teachers, develop a valid communicative assessment program for a full school year (a compilation of formative and summative communicative tests) that could easily be implemented in Dutch low-form education and that would incite positive washback. The scientific objective was to further identify and specify the practical challenges (regarding feasibility) and the conceptual challenges (regarding validity) teachers face while designing this communicative assessment program, and to formulate suggestions to overcome them. As such, the research question guiding the current study is: which challenges regarding feasibility and validity do FL teachers face in the design of communicative test materials, and what decisions can be taken to overcome them?

In this paper we present the first results of this co-design study.

Method

Participants

A group of 21 FL teachers of the most commonly taught foreign languages in Dutch secondary schools (English, French, German, and Spanish) from 15 different schools took part in a Professional Learning Community (PLC) with the aim of learning by design, taking both theoretical and practical aspects into consideration in the selection and (re)design process of a communicative assessment program. The participants were all graduated teachers with varied years of experience ranging from 1 to 25 years. They all taught at least one class in lower form.

Participants of the PLC took part in eight interactive working-sessions of 3 hours each. The sessions were organized and supervised by the researcher (i.e., the first author). During the eight sessions, participants shared their vision on communicative teaching and testing practices, compared their practice with theory on communicative testing, shared their testing materials (as used in their current practices), (re)designed tests, and compiled tests to create a feasible communicative assessment program.

Data collection and analysis

Notes on the proceedings of each PLC session as well as the exchanges, remarks, and questioning of the participants were reported by the researcher in a logbook. The report of each session was submitted to each of the participants individually for a member check to ensure transparency. They could complete, modify and/or confirm the report.

The data were analyzed inductively, according to the steps specific to qualitative research analysis (Corbin & Strauss, 2008). In this case, the codes were short sequences of words that described the challenges participants faced during the design process and the decisions they made to overcome them. Codes were then synthesized and categorized into two themes: practical challenges and decisions (regarding feasibility) and conceptual challenges and decisions (regarding validity).

First results

Challenges and design decisions regarding feasibility

Three challenges regarding feasibility were identified:

1. Lack of materials available for learning activities.
2. Lack of time allocated for test administration and scoring.
3. Students' lack of experience with communicative tests.

In response to each of these concerns, decisions were made within the PLC to enhance feasibility in the design of the test materials. The different suggestions to overcome these challenges are reported below.

Lack of materials available for learning activities

Results from the logbook show that the majority of the participants used a textbook and did not have much time to develop their own learning activities in addition to it. Their first concern in the design of communicative tests was to be able to keep using the learning activities from their textbooks while preparing their students for the tests. Most of the learning activities and tests from the textbooks used are organized thematically (e.g., sport, school, holidays) and focus mainly on knowledge of vocabulary, chunks and grammar rules out of context. Communicative learning activities are included, but do not constitute the main focus.

To overcome this problem, we decided to develop communicative tests that can be adapted to the different themes covered by the textbooks. Test tasks focus on language acts to be performed (such as persuading, describing, giving an opinion, etc.) in accordance to the corresponding CEFR level, and can be adapted to any theme. In this way, all activities about a specific theme in the textbook can still be used in preparation of the test, and teachers only need to put more emphasis on the communicative activities in a specific chapter. This implies that teachers use their textbooks as a tool and no longer as the curriculum as such. It is still crucial to select a textbook that provides enough opportunities for communicative activities.

Lack of time allocated for test administration and scoring

The second concern teachers expressed in the PLC meetings is the time allocated to teachers to administrate and score tests. Communicative testing requires more time. Writing, reporting, and speaking are complex operations, and the assessment of these skills requires a lot of attention and expertise from the teachers. One solution to compensate for the time-consuming nature of communicative testing is simply to test less often and introduce more formative activities in between less frequent summative assessments. Another challenge resulting from this decision was to be able to test the four language skills at the intended level, without creating pressure at the end of the year. To resolve this problem, we chose to test the different skills in an integrated way, in accordance with the communicative approach.

The time spent on administering a communicative test in classes of about 30 students was another concern, particularly regarding the assessment of speaking and conversational skills. We suggested using teaching time to assess speaking skills by organizing and scoring presentations during the lessons. This can be justified by the fact that, while the presentation is a test for some students, it can function simultaneously as a learning activity for the others. Besides, presenting in front of an audience as in real life enhances authenticity. To assess conversational skills, we designed test tasks that could be performed by three students at the same time. Per group students receive the same realistic situation in which each of them has to perform a different task.

Finally, teachers were also concerned about the time needed to score communicative tests. We decided therefore to develop rating scales with a holistic part in addition to an analytical part. The holistic part is short and to the point, and can be scored very

quickly to give a first indication of the global CEFR level at which a student performed a communicative act. The analytical part zooms in on the quality of communicative competence. The criteria in the analytical part of the scale are descriptive, based on CEFR Can Do statements. The details of the descriptions allow teachers to score quickly, without having to give extra feedback to justify the score.

Students' lack of experience with communicative tests

The last practical challenge mentioned by the teachers was to keep students motivated with fewer grades and tests that do not focus on reproduction only. Tests assessing FL skills are often perceived by students as being less focused and therefore more difficult to prepare for.

To prevent students from 'hiding' and not working regularly or effectively between the formal assessment moments, we suggested the systematic introduction of formative communicative activities during the lessons.

Challenges and design decisions regarding validity

Three challenges regarding the validity of communicative tests were identified.

1. How to introduce unpredictability.
2. How to integrate grammar and vocabulary.
3. How to integrate language skills.

To ensure a higher degree of 'communicative validity' in the design of test materials (in order to stimulate positive washback), we made decisions within the PLC to overcome these challenges.

Introduction of unpredictability

Teachers were concerned by the introduction of unpredictability to enhance validity, because they did not know how to prepare their students to react spontaneously and adequately to unprepared questions and situations. To overcome this challenge, we decided in the design of the test tasks to focus on students' ability to perform speech acts that are present in all kinds of realistic situations (reporting, corresponding, sharing experience and information, expressing opinions, etc.). Themes coming from their textbook (sport, holidays, school, etc.) will be known in advance, but students will have to perform one of the acts within a new situation. This increases unpredictability while still providing sufficient guidance to prepare students.

Integration of grammar and vocabulary

A large number of the participating teachers were still testing knowledge of grammar and vocabulary and chunks out of context. It was challenging for them to discard this type of testing and replace it with the assessment of language skills alone. They were afraid that students would no longer learn enough vocabulary, chunks, and grammar rules if these were not tested directly. We decided to create two categories in the rating scales to assess the use of vocabulary and grammar in the context of a communicative task. In addition, we insisted on the importance of using formative activities addressing the development of knowledge of vocabulary and grammar rules next to communicative learning activities.

Integration of language skills

The last challenge faced by the teachers was the integration of language skills within a test to enhance authenticity. Different combinations were possible for this integration. We chose to integrate reading and writing skills on the one hand and the listening and speaking skills on the other, with a distinction between conversational skill (listening and speaking in the context of a conversation) and listening and speaking skills (listening to gather information and then presenting something about it). This combination appears to be the most common in real life.

Teachers were concerned by the fact that one of the skills could be underrepresented in a test or that students could compensate one skill with another, which would not give a clear idea of the level of each skill independently. In order to enhance validity on that point, we introduced Can Do statements in the analytic part of the rating scales belonging equally to each language skill.

Conclusion

This paper presented the first results regarding practical and conceptual challenges faced by FL teachers in the design of communicative test materials and the decisions made to overcome them.

Practical challenges often related to the limited time allocated to teachers to prepare lessons, and to develop or select communicative learning/assessment activities. This time issue appeared to be even more pressing when the textbook used did not include sufficient communicative learning activities. Other practical challenges were related to tests administration and scoring within the allocated time. Teachers were also concerned by students' motivation, with fewer tests focusing only on language skills.

In addition to these practical challenges, teachers also faced conceptual challenges regarding the operationalization of unpredictability, the integration of grammar and vocabulary within a communicative task, and the integration of different language skills within one assessment.

Decisions made to address these challenges are summarized below in the form of suggestions to be taken into account when designing test materials for communicative learning goals:

1. Language skills should be tested in an integrative way to reduce the amount of tests and enhance authenticity.
2. Test tasks should include new situations to ensure unpredictability, but should be formulated with precise instructions to provide enough guidance to prepare students.
3. Rating scales should assess language skills in an integrated way. They should be descriptive, based on Can Do statements belonging equally to each language skill. They should include a holistic part and an analytic part to save time in scoring and to enable efficient feedback.
4. Formative activities should be systematically introduced to keep students motivated and to guide their learning process.

References

- Alderson, C., & Wall, D. (1993). Does washback exist? *Applied Linguistics*, 14(2), 115–129.
- Biggs, J. (1996). Enhancing teaching through constructive alignment. *Higher Education*, 32(3), 347–364.
- Brown, J. D. (2005). *Testing in Language Programs*. New York: McGraw-Hill.
- College voor Toetsen en Examens. (2017). *Examenprogramma taal vanaf CE 2017*. Retrieved from : www.examenblad.nl/
- Corbin, J., & Strauss, A. (Ed.). (2008). *Basics of Qualitative Research* (3rd edn). Thousand Oaks: Sage.
- Fasoglio, D., de Jong, K., Pennewaard, L., Trimboos, B., & Tuin, D. (2015). *Moderne vreemde talen: Vakspecifieke trendanalyse 2015*. Enschede: SLO.
- Fulcher, G. (2003). *Testing Second Language Speaking*. London: Pearson Longman.
- Green, A. (2007). Washback to learning outcomes: a comparative study of IELTS preparation and university professional language courses. *Assessment in Education*, 14(1), 75–97.
- Hakim, L. N. (2018). Washback effect in language testing: What do we know and what is its effect?. *Jurnal Forum Didaktik*, 2(1), 59–68.
- Harding, L. (2014). Communicative language testing: Current issues and future research. *Language Assessment Quarterly*, 11(2), 186–197.
- Kissau, S., Rodgers, M., & Haudeck, H. (2015). Practicing what they preach ? A comparison of teacher candidate beliefs and practices. *International Journal of Language Studies*, 9(4), 29–54.
- Larsen-Freeman, D., & Anderson, M. (2011). *Techniques and Principles in Language Teaching*. Oxford: Oxford University Press.
- Morrow, C. K. (2018). Communicative language testing. In J. I. Lontas (Ed.), *The TESOL Encyclopedia of English Language Teaching* (pp. 342–350). New Jersey: John Wiley & Sons Inc.
- Schnabel, P., Ten Dam, G., Douma, T., van Eijk, R., Tabarki, F., Van der Touw, A., & Visser, M. (2016). *Ons onderwijs 2032: Eindadvies*. Retrieved from: www.rijksoverheid.nl/documenten/rapporten/2016/01/23/eindadvies-platform-onderwijs2032-ons-onderwijs2032
- West, L., & Verspoor, M. (2016). An impression of foreign language teaching approaches in the Netherlands. *Levende Talen Tijdschrift*, 17(4), 26–36.